NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

FRAME RATE EFFECTS ON HUMAN SPATIAL PERCEPTION IN VIDEO INTELLIGENCE

by

Kurt A. Kempster

September 2000

Thesis Advisor:	Rudolph P. Darken
Associate Advisor:	Terrance C. Brady

Approved for public release; distribution is unlimited

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2000	**3. REPORT TYPE AND DATES COVERED** Master's Thesis September 1998-Sepetember 2000

4. TITLE AND SUBTITLE: Title (Mix case letters) Frame Rate Effects on Human Spatial Perception in Video Intelligence	**5. FUNDING NUMBERS**
6. AUTHOR(S) Kurt A. Kempster, Major USMC	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)*

This thesis examines the effect that the frame rate of a streaming video feed has on one's ability to maintain spatial perception. It defines the current technologies available to capture and encode digital video. It describes the current and near future wireless information systems that could be utilized to support streaming video.

This thesis investigates through experimental trials of subjects viewing video streams at different frame rates, the effect those frame rates have on the subject's spatial perception. This thesis analyzes and summarizes the data collected from this experiment and provides recommendations. It is determined that the inherent chaotic nature of tactical movement and the method used to encode digital video are not compatible for video streams with high motion in the three dimensional planes. Results of this analysis suggest that a large amount of bandwidth would be consumed to provide the minimum quality of service indicated by the data and suggests that video to the commanders at the frontline is not a useful allocation of bandwidth.

14. SUBJECT TERMS Reconnaissance, Digital Video Imagery, Real-time Video, Wireless Communications, Bandwidth, Spatial Perception	**15. NUMBER OF PAGES** 102
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

FRAME RATE EFFECTS ON HUMAN SPATIAL PERCEPTION IN VIDEO INTELLIGENCE

Kurt A. Kempster
Major, United States Marine Corps
B.A., Rollins College, 1985

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
September 2000

Author: Kurt A. Kempster

Approved by: Rudolph P. Darken, Thesis Advisor

Terrance C. Brady, Associate Advisor

Dan C. Boger, Chairman
Information Systems Academic Group

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis examines the effect that the frame rate of a streaming video feed has on one's ability to maintain spatial perception. It defines the current technologies available to capture and encode digital video. It describes the current and near future wireless information systems that could be utilized to support streaming video.

This thesis investigates through experimental trials of subjects viewing video streams at different frame rates, the effect those frame rates have on the subject's spatial perception. This thesis analyzes and summarizes the data collected from this experiment and provides recommendations. It is determined that the inherent chaotic nature of tactical movement and the method used to encode digital video are not compatible for video streams with high motion in the three dimensional planes. Results of this analysis suggest that a large amount of bandwidth would be consumed to provide the minimum quality of service indicated by the data and suggests that video to the commanders at the frontline is not a useful allocation of bandwidth.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYM

AM	Amplitude Modulation
ATM	Asychronous Transfer Mode
CBR	Constant Bit Rate
CODEC	Code/Decode
DCT	Discrete Cosine Transform
DoD	Department of Defense
EHF	Extremely High Frequency
FM	Frequency Modulation
FM	Field Manual
FSK	Frequency Shift Keying
GOP	Group of Pictures
HDTV	High Definition Television
HTTP	HyperText Tranfer Protocol
IP	Internet Protocol
ITU	International Telecommunications Union
JV	Joint Vision
JCS	Joint Chiefs Of Staff
JPEG	Joint Photographic Experts Group
KBPS	Kilo Bits Per Second
LAN	Local Area Network
LDR/MDR	Low Data Rate/Medium Data Rate

MCDP	Marine Corps Doctrinal
MHz	MegaHertz
MMX	MultiMedia Extension
MPEG	Motion Pictures Experts Group Publication
NPS	Naval Postgraduate School
NTSC	National Televison Systems Committee
OODA	Observe, Orient, Decide, Act
PAL	Phase Alternating Line
PC	Personal Computer
POTS	Plain Old Telephone Service
QoS	Quality of Service
RAM	Random Access Memory
RF	Radio Frequency
RGB	Red Green Blue
RLE	Run Length Encoding
RSS	Regression Sum Squares
SATCOM	Satellite Communications
SECAM	Sequential Coleur Avec Memoire
SVT	Scalable Video Technology
TV	Television
UDP	User Datagram Protocol
USAF	United States Air Force

USMC	United States Marine Corps
VCR	Video Cassette Recorder
WYSIWYG	What You See Is What You Get

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to extend my sincerest thanks to Professor Rudolph P. Darken for the guidance and support he has given me throughout this research process. I would also like to thank the forty plus volunteers who took time from their busy schedules to act as pilot and actual subjects for my research experiment.

Most importantly I would like to thank my wife, Melissa Kempster, for her unwavering support and understanding as I went through this process. She has always been and will always be my biggest supporter.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

The success or failure of a unit in battle relies heavily on the decisions of the unit commander and subordinate leaders. Information superiority, not only in terms of quantity but also of quality, is a critical element of this aspect of warfare. The desire to increase the ability of the commander to shorten his decision-making process is a driving force in the digitization of the battlefield. However, the decision-making process is most often constrained by the ability to gather, disseminate, and comprehend what can be referred to as useful information such that it can be used effectively

The quicker a commander can make informed decisions, as compared to the enemy commander, the greater his ability to achieve objectives on the battlefield. The commander who makes good decisions and executes these decisions at a superior tempo in the face of uncertainty and constrained time most often leads his forces to victory. Munroe/Pasagian (1998) showed that it is possible to capture video images and then inject them into an information network for the commander to consider. Many things affect human decision-making processes to include stress and time compression. As time is compressed and stress increases, decision makers may; rely on a limited fraction of the available information; concentrate more on decisions based on an obsolete understanding of the environment and less on situational awareness; and increase their micro-management of subordinates (Munroe/Pasagian 1998).

Current Marine Corps doctrine calls for the use of verbal and pencil sketch data in reconnaissance missions (Field Manual 5-170 [FM 5-170], 1998). Not only does this type of data tend to be error prone, but also it is also time-consuming to collect and use. If reconnaissance data could be collected from the field in its natural form (imagery or text) and in a timely fashion, this would greatly increase the tempo of the commander's decision-making process. However, it is not safe to assume that more information is always better. Information saturation can be a continual, real-life problem. A reconnaissance doctrine that called for widespread use of digital video imagery, voice, and textual data all streaming to the commander simultaneously would likely prove useless at best and harmful at worst. (Munroe/Pasagian 1998)

The military is aggressively pursuing integration of technology into the command and control process to take advantage of the rapid pace of change with regard to information technology in the armed forces. Every day we are awed by new developments in science and technology and the military opportunities and threats they represent. The use of digital video images, as information is an example of such a development.

Delivery of superior information to the commander in the form of imagery is central to the research contained within this thesis. Enhanced information delivery for the purpose of improving a commander's ability to speed decision making, requires a close examination of the decision making cycle.

1. Boyd's Theory

For crisis decision making, Boyd's cycle, developed by USAF Col. John Boyd, is one of the most useful models of the decision making process. Boyd's cycle was developed with adversaries and opposing wills in mind. However, it can be applied in other crisis situations as well. Boyd's cycle describes conflict in a time-competitive environment, which is cyclic in nature. Two opposing wills present a series of unexpected and threatening situations to one another. The side that cannot keep pace with the threatening situations is defeated. This happens regardless of the size, strength, or equipment possessed by the forces. Boyd's cycle has four distinctive phases, observation, orientation, decision, and action. Together they complete one cycle. Boyd's cycle is also known as the OODA loop (Marine Corps Doctrinal Publication 6 [MCDP 6], 1996).

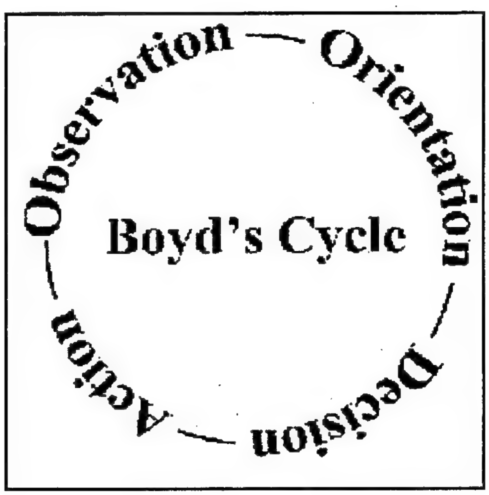


Figure 1.1 Boyd's Cycle.

The first phase is observation. Observation refers to the necessity of becoming aware, especially through careful and directed attention. The decision-maker must observe what is taking place and determine the circumstances under which he or she must function. Observation always involves one or more of the five senses. Sometimes we seek information and sometimes it is thrust upon us.

The second phase is orientation. Orientation is described as the state of locating or placing an item in relation to something else. Orientation is distinct from observation since this is when the initial assessment begins and some type of prioritization is necessary. Orientation is a synopsis or summary of the previous observation that helps bridge observations to the decisions they influence. It is a mental "snapshot" of the incident. This is required because the situation is too fluid and changing to make a sound decision without making it static, even if only for an instant.

The third step in Boyd's cycle is decision. Decision refers to the passing of judgment on an issue under consideration. This is the step in which a commander attempts to control a situation in which he finds himself. This determines what the commander's next course of action will be. Decision converts the information into orders. Based on orientation we make a decision, either an immediate reaction or a deliberate plan.

The last step in the OODA loop is action. Action refers to the state or process of acting or doing something. Action is where the decision is put into effect. The OODA loop is continuous; as you act you observe the results, the process starts all over again. It is possible, and very probable; to have multiple OODA loops, in various stages, spinning

at the same time, but not necessarily at the same rate. The OODA loop reflects how command and control is a continuous, cyclic process.

The goal of integrating technology into the command and control process is an increased operational tempo in order to seize the initiative and overwhelm one's enemy by being able to observe, orient, decide, and act (OODA) faster than he is able to. "Speed is an essential element of effective command and control. It means shortening the time needed to make decisions, plan, coordinate, and communicate" (MCDP 6, 1996).

2. Current Philosophies and Doctrine

The trend toward the integration of new technological tools must be conducted carefully. Standardization of equipment, interoperability and associated relevant issues must be considered. There are two existing documents that present ideas for the application of technology as a tool and have been the genesis of the push to thrust technology into all aspects of the war fighting process.

A common direction for each of the Armed Services is developed within Joint Vision 2020 (JV 2020). Since leveraging technological opportunities is central to JV 2020, it is necessary to consider the concepts put forth by the Joint Chiefs of Staff (JCS).

Further, the Marine Corps Doctrinal Publication 6 (MCDP 6) is addressed within this document in order to portray central themes of command and control theory and philosophy in the Marine Corps:

> That command and control is not the exclusive province of senior commanders and staffs; effective command and control is the responsibility of all Marines (MCDP 6, 1996).

a) Joint Vision 2020

Joint Vision 2020 (JV 2020) seeks to form a template for how our Armed Forces will prepare to fight and operate into the 21st century. The JCS plans to achieve dominance through JV 2020 by recognizing that the future of warfighting is embodied in improved intelligence and command and control (Joint Vision 2020 [JV 2020], 2000). Historically, technology embodies the tools that leaders and managers seek in order to manipulate a situation to produce favorable results. More than ever before, a command and control system is crucial to success on the battlefield and must support shorter decision cycles and instantaneous flexibility in an operational environment.

In preparing for the 21st century, Joint Vision 2020 develops four important operational concepts integral to the Armed Forces ability to dominate an adversary. These are (1) dominant maneuver, (2) precision engagement, (3) full dimensional protection and (4) focused logistics.

Of the four operational concepts put forth by Joint Vision 2020, those of dominant maneuver and precision engagement are central to the information superiority that may be achieved through the delivery of real time video imagery. Both concepts allow our Armed Forces to gain a decisive edge through responsive command and control. Dominant maneuver allows forces to gain an advantage by controlling each aspect of the battle space (JV 2020, 2000). This is accomplished through a combination of decisive speed and tempo. Both speed and tempo in maneuver are achieved through the employment of improved sensors and real-time evaluation.

Precision engagement also allows forces to gain an advantage by shaping the battle space. This is accomplished through high fidelity target acquisition, prioritized target requirements and accurate weapons delivery techniques (JV 2020, 2000).

b) Marine Corps Doctrinal Publication 6 (MCDP 6)

According to the Commandant of the Marine Corps (CMC), the Marine Corps' view of command and control is based on the common understanding of the nature of war and the Corps' warfighting philosophy. It accounts for the timeless attributes of war, as well as the impacting features of the information explosion, resulting from modern technology. MCDP 6 addresses the complex environment of command and control *(uncertainty and time)* and theory of command and *control (to include the OODA loop, image theory, and decision-making theory).*

The operational environment is characterized by a dynamic, fluid situation. In such a chaotic setting, commanders and staffs must tolerate ambiguity and uncertainty, identify patterns, seek and select critical information and make rapid decisions under stress (MCDP 6, 1996). Command and control systems must therefore be planned as extensions of the human senses and processes to help commanders *reduce uncertainty*, form perceptions, react, and make *timely decisions*. This allows commanders to be effective during high-tempo operations.

People assimilate information more quickly and effectively as visual images than in text. We can say that an image is the embodiment of our understanding of a given situation or condition (Zimm, 1999). For these reasons, a commander can ensure

a more agile and decisive response to his environment than his enemy-- and that means victory on the battlefield.

3. Issues Related to Video Transmission

Video to the commander has become the catch phrase as the technology in terms of availability and reliability has increased. There are a multitude of service providers offering web hosting and video to the desktop using IP multicast and other streaming media protocols. However these applications have the benefit of robust ground stations and high-speed network connections.

Hollywood has put the notion in observer's heads that we can stream video from a Marine in the field to the commander back at any location. In movies such as *The Rock* and *Alien 2* they portray a television quality transmission from a wireless transmission routed back to the commander. The technology to transmit video to the commander does exist but not at the television quality Hollywood would like us to believe. This notion has permeated the culture and it is the expected quality of service from real systems. <u>The fundamental question that underlies the transmission of video at any level of quality of service is whether or not the person viewing it can derive useful information from it. For the purpose of this thesis, we will define "useful information" as spatial perception; Can the viewer indicate where he is in the observed environment on a map of the same area?</u>

a) Quality of Service (QoS)

The quality of service is the key aspect of integrating video into the command and control system. How useful a video stream is to the commander is directly

related to the type of video being transmitted, at what rate in terms of bits per seconds and frames per second and the method in which it is transmitted. Some of these things can be controlled, such compression algorithm/type and data rate, while others, amount of motion in the actual video source and subject can not in a tactical environment. All of these things all affect the quality of the transmission. Through minimizing the number of variables such as content and type of the video we will define quality of service as the point where useful information can no longer be extracted form a particular stream. This will lead to the determination of how QoS affects the usefulness of the video to the observer.

B. OBJECTIVES

This thesis is narrowly focused on researching what effects various frame rates and their resultant quality of service has on the end user to maintain one's spatial awareness from viewing streaming video from a unit in the field. The goal is to determine, (1) Is it even possible to maintain spatial perception while observing streaming video and (2) If it is possible, at what level of quality of service can a commander remain oriented in a video feed. The end result will be a determination of the benefits of streaming video to the commander, architecture requirements required to attain these rates and supportability of these rates considering existing and planned near term systems. The the intent of this thesis is not to determine an actual system configuration to support video from the field, but to seek out the effects quality service has on utility to the user and if there is utility to the user the level of performance to maintain that utility.

The objective is to determine the requirements in terms of frame rate, bandwidth and supportability for delivering real-time imagery from forward deployed reconnaissance units to the commander in the rear, thus enhancing the commander's decision making capabilities.

This thesis examines the following research questions:

1. What are the frame rates that are associated with current and proposed future technologies that may be used for video intelligence?

2. Are viewers of streaming video at these frame rates able to maintain spatial awareness while viewing?

3. What effect if any does frame rate/fidelity of streaming video via wireless systems have on the user's ability to maintain spatial perception?

4. How does the video enhance the user's ability to make decisions in the high tempo environment of ground combat in an urban environment?

5. What level of video fidelity is needed in order to achieve realistic, credible and effective aids to the decision-making process for a ground commander?

C. ASSUMPTIONS

With the advent of high-speed Internet access the proliferation of commercial based "web casts" of video content is exploding. The civilian sector is responding through a variety of hosting services coming into the marketplace. These services are not designed to be used in the expeditious environment one would find on the battlefield of

the future. Traditionally, the bandwidth limitations associated with most networks have made the transmission of video cumbersome, impractical, expensive, and of poor quality. One response to this has been to make the bandwidth bigger, known as broadband. The broadband push of the information technology industry is largely fueled by this increasing demand for these wide varieties of video applications. The other solution to the bandwidth restrictions of networks is to develop more efficient methods for encoding of the video signal to ensure from source to receiver.

The ability to capture video for transmission over any network is generally a routine task. Digital video cameras are available on the marketplace and they can output the video in many of the existing video standards, MPEG-1, -2, and –4. One example of this is the Sharp Corporation Model VN-EZ1U MPEG-4 Digital Recorder. It can transmit images at rates as low as 28.8 kpbs. (Sharp 1999)

The link from the camera to the uplink device is technologically available and will not be discussed in this thesis. Further, in-depth analysis in the communication, encryption and associated technologies will not be addressed.

D. METHODOLOGY

The following methodology was used in the preparation of this thesis:

1. Background and analysis of the physiology of human spatial perception.
2. Research of the current video compression and standards
3. Examination of satellite based information systems available now and in the near future to determine theoretical bit rates available to support deep reconnaissance.
4. Development of a prototype commander's station for streaming video.

5. Conduct of an experiment to determine at what level of QOS does streaming video become a hindrance vice enhancement to the decision making process of the commander in the scope of the prototype commander's station.

E. ORGANIZATION OF THE THESIS

Chapter II provides background information on the types of video compression algorithms and the QOS associated with each one. It defines the bandwidth requirements for each method and the benefits/drawbacks to each.

Chapter III describes the existing and near future satellite communication assets that could be utilized to stream video back to the commander. It will examine the bit rates, bandwidth supported and system characteristics needed to complete the link at the required throughput if possible.

Chapter IV provides an overview of how a human spatial perceives things and discusses the experiment methodology and results.

Chapter V recommends an estimate of supportability for the information architectures in support of the resultant frame rates from the experiment. It will also cover any recommendation or suggestions for further study to include processes and equipment.

II. VIDEO STANDARDS AND FORMATS

It is important to understand the processes that go into creating the content that one might want to stream from a source to a user. It is not simply plugging a camera into a transmitter and beaming the picture to the user.

A. VIDEO BASICS

As we examine the types of video that are available, it is useful to pause and highlight some basics of video transmissions and display. No attempt will be made to explain every intricacy of video just those that suffice as background and are applicable to this thesis subject. The quality of the video transmission is dependent on a variety of variables: bandwidth, type, analog or digital, refresh rates and synchronization. Video, as it is often thought of as a VCR tape or a broadcast on your television set is analog. A video is drawn from left to right, top to bottom. Each scan is a single horizontal pass across the screen. This is followed by a horizontal blanking interval, where the electron gun moves to the beginning of the next scan line. After every scan line has been drawn, a vertical blanking interval allows the electron gun to move from the lower right to the upper left, and the process begins again.

1. Analog Video Formats

There are many standards that govern the creation and transmission of video content around the world. This fact alone can cause there to be inoperability problems

across global markets. To highlight the various formats the standards organizations are listed below:

- NTSC (National Television Systems Committee) is the standard broadcasting system for North America, Japan, and a few other countries. It has 525 lines of resolution with a 30Hz frequency rate.
- SECAM (Sequential Coleur Avec Memoire) is a frequency-modulated signal that has 625 lines of resolution and a 25Hz refresh rate. It is used in France and Eastern Europe.
- PAL (Phase Alternating Line) is similar to SECAM and is used in parts of Western Europe.

Because human vision is wider than it is tall, television and video displays are rectangular, with the width being greater than the height. The ratio of the width to the height is called the aspect ratio. For standard TV i.e. NTSC, SECAM, and PAL the aspect ratio is 4:3, giving a resolution of 700x525 in the case of NTSC.

A number of activities aimed at setting new High-Definition television (HDTV) standards are taking place worldwide. Common to the HDTV standards are a widened aspect ratio (16:9 vice 4:3, increased picture resolution, and audio of compact disc quality. North America has taken the approach of formulating a fully digital HDTV standard. The new HDTV standard has 1000 lines of resolution and an aspect ratio of 16:9, giving resolution of 1778x1000. [Ragahavan 1997]

2. The H.261 Standard (*p x 64*)

The H.261 standard, commonly called *p x 64,* is optimized to achieve very high compression ratios for full color, real time motion video transmission. The p x 64 compression algorithm combines intraframe and interframe coding to provide fast processing for on the fly video compression and decompression. The standard is optimized for video-based telecommunications. Because these applications tend not to be motion-intensive, the algorithm uses limited motion search and estimation to achieve higher compression ratios. For standard video communication images compression ranges are from 100 to 2000:1. [Laplante 1996]

3. The H.263 standard

The H.263 standard, published by the International Telecommunications Union (ITU), supports video compression (coding) for video-conferencing and video-telephony applications at very low bit rates.

a) *Applications*

- Videoconferencing and video telephony have a wide range of applications including:
- Desktop and room-based conferencing
- Video over the Internet and over telephone lines
- Surveillance and monitoring
- Telemedicine (medical consultation and diagnosis at a distance)
- Computer-based training and education

In each case video information (and perhaps audio as well) is transmitted over telecommunications links, including networks, telephone lines, ISDN and radio. Video has a high "bandwidth" (i.e. many bytes of information per second) and so these applications require video compression or video coding technology to reduce the bandwidth before transmission. (ITU-T, 1999)

4. Image Concepts and Structures

According to trichromatic theory, the sensation of color is produce by selectively exciting three classes of receptors in the eye. In an RGB system, color is produced by combining the three primary colors: red, blue, and green (RGB). Another representation of color images better suited to the compression of images is the YUV representation. YUV describes the luminance and chrominance of the image. Luminance (Y), which provides the gray-scale version of the image, and Chrominance (U) and Chrominance (V) that convert the gray-scale image to a color image. This is more natural for image compression and is used intensively. [Rao 1996]

5. Refresh Rates

The human eye can distinguish movement at about 1/16 of a second. Despite this, some flicker can be seen even at 30 frames per second. In order for the human eye not to perceive flicker in a bright image, the refresh rate of the image must be higher than 50 frames per second. However, to speed up the frame rate to that rate while transmitting the whole frame data would require speeding up the scanning, both vertical and horizontal, thereby increasing the bandwidth. In order to alleviate this problem, interlacing is used.

Interlacing draws odd scan lines in the first 1/60 of a second and then draws the even scan lines in the second 1/60 of a second, effectively converting a 30 Hz signal to a 60Hz refresh rate while keeping the same bandwidth as the original signal.

6. Synchronization

An interesting problem is inherent in NTSC video. The advertised frame rate of NTSC video is 30Hz, however due to a harmonic interference with the color carrier; the frame rate was dropped to 29.97Hz a 0.1% decrease in the frame rate. Because synchronization information is represented as hh:mm:ss:ff (hour: minute: second: frame#), this poses a serious synchronization problem. If we assume each frame is 1/30th of a second, then display time will drift away from the presentation time. This problem can be overcome by dropping the first two frame numbers, not the actual frames, of every minute divisible by ten.

B. MAKING DIGITAL MEDIA FROM ANALOG MEDIA

The most commonly used video cameras take an analog sample and must convert to a signal before transmission over a digital network. There has been a recent influx of video cameras that record in a digital format but the cost of these cameras is currently cost prohibitive ($900-$1200) and they are not durable enough for field use.

The bandwidth required for digital video is staggering. Uncompressed NTSC video requires a bandwidth of 20Mbyte/sec, HDTV requires 200Mbyte/sec. Various encoding techniques have been developed in order to make digital video feasible. Two

classes of encoding techniques are Source Encoding and Entropy Encoding. I will discuss both but will focus on the coding techniques used in H.263.

1. Source Encoding

Source encoding is lossy and applies techniques based upon properties of the media. There are four types of source encoding:

- Sub-band coding gives different resolutions to different bands. E.g. since the human eye is more sensitive to the intensity changes than color changes, we separate the video signal into different components like Y, U and V components. Sub-band coding facilitates subsampling.
- SubSampling groups pixels together into a meta-region and encodes a single value for the entire region.
- Predictive coding uses one sample to guess the next. It assumes a model and sends only the differences from the model (error values).
- Transform encoding transforms one set of reference planes to another.

2. Entropy Encoding

Entropy encoding techniques are lossless techniques that tend to be simpler than source encoding techniques. The three entropy encoding techniques are:

- Run-Length Encoding (RLE) encodes multiple appearances of the same value as {value, # of appearances}. E.g.1, 1,1,1,2,2,2,3 would encode as {1,4}, {2,3}, {3,1}.

- Huffman Coding looks at statistical distributions of data to provide compression. It does this by giving the smallest length code to the most frequent character and then giving the longest code to the character that occurs least. With Huffman coding, any code cannot be a proper prefix of another code. If this property did not hold, we would be unable to decode the variable bit-length code, because one value could appear as a combination of two other values or vice versa.
- Arithmetic coding is similar to Huffman coding, but is more complex and provides better compression, especially for text. For images it is not necessary. [Ragahavan 1997]

The entropy encoding in H.263 is based on the Huffman technique and is used to compress the quantized DCT coefficients. The result is a sequence of variable-length binary codes. These codes are combined with synchronization and control information (such as the motion "vectors" required to reconstruct the motion-compensated reference frame) to form the encoded H.263 bit stream.

3. Video coding in H. 263

A typical system is shown in Figure 2.1.

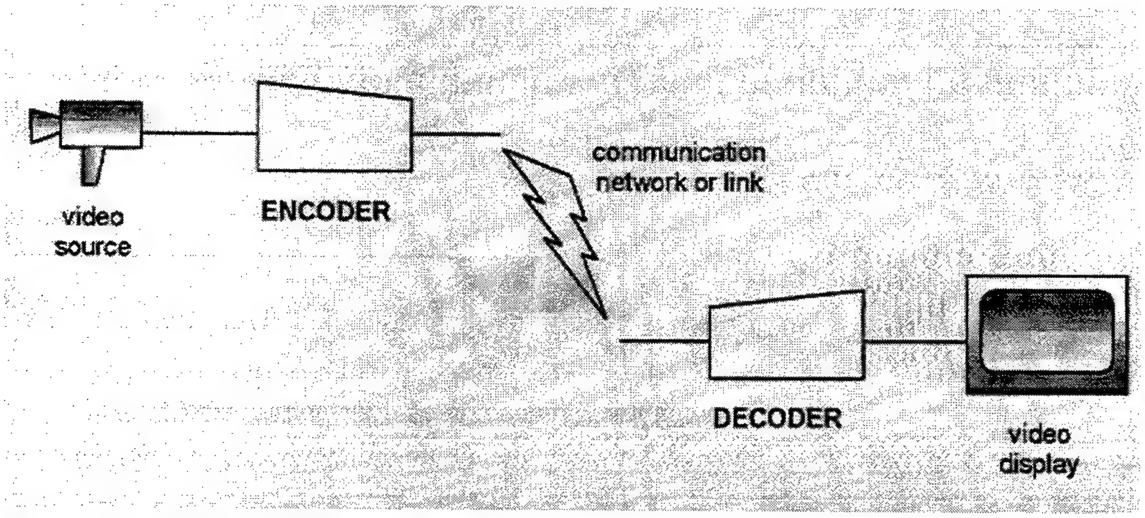


Figure 2.1 H.263 Video System

Frames of video information are captured at the source and are encoded (compressed) by a video encoder. The compressed "stream" is transmitted across a network or telecommunications link and decoded (decompressed) by a video decoder. The decoded frames can then be displayed. (4I2I, 2000)

a) The H.263 System

A number of video coding standards exist, each of which is designed for a particular type of application: for example, JPEG for still images, MPEG2 for digital television and H.261 for ISDN video conferencing, as discussed earlier. H.263 is aimed particularly at video coding for low bit rates (typically 20-30kbps and above). The H.263 standard specifies the requirements for a video encoder and decoder. It does not describe the encoder or decoder itself: instead, it specifies the format and content of the encoded (compressed) stream. A typical encoder and decoder are described here. Many of the

details of the H.263 standard have been "skipped" such as syntax and coding modes because they do not fall into the scope of this work. (4I2I, 2000)

b) H.263 Encoder

The below is a sample H.263 encoder. The details of the diagram will be discussed in the following section.

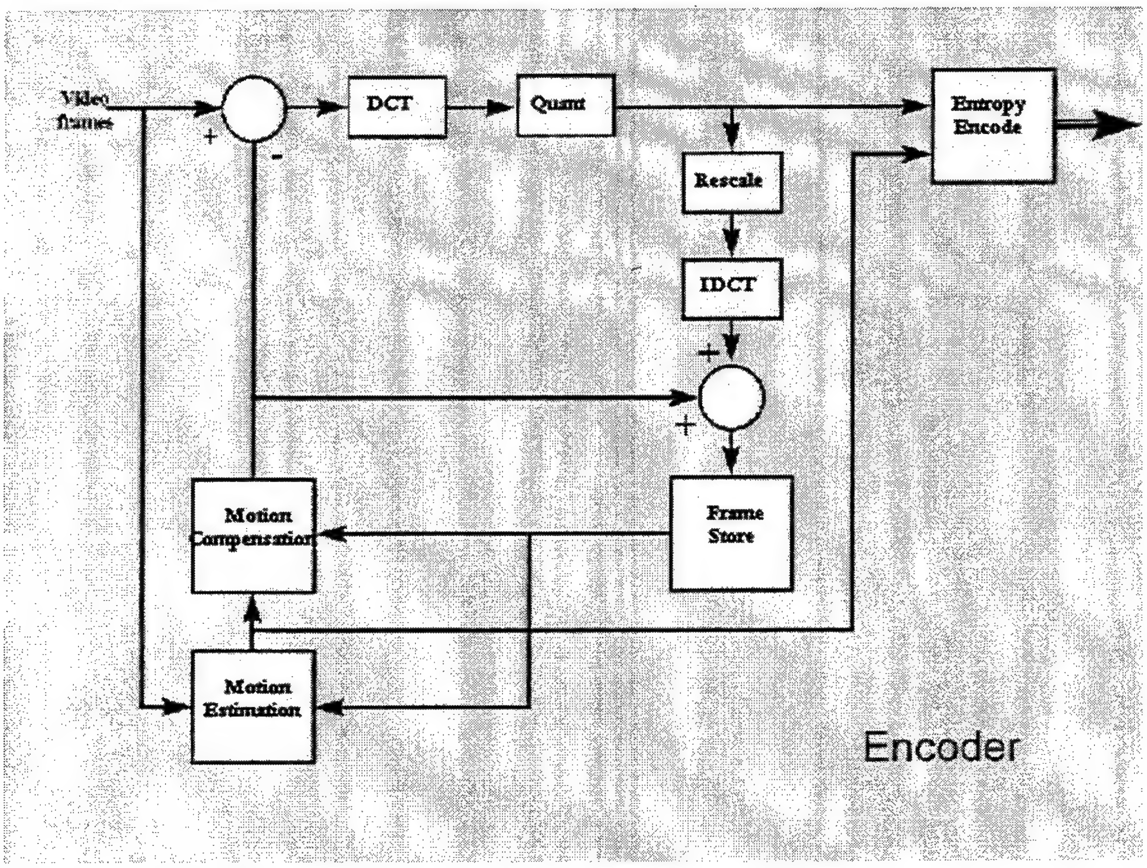


Figure 2.2 H.263 Encoder

c) Motion Estimation and Compensation in H.263

The first step in reducing the bandwidth is to subtract the previous transmitted frame from the current frame so that only the difference or residue needs to be

encoded and transmitted. This means that areas of the frame that do not change (for example the background) are not encoded. Further reduction is achieved by attempting to estimate where areas of the previous frame have moved to in the current frame (motion estimation) and compensating for this movement (motion compensation). The motion estimation module compares each 16x16 pixel block (macroblock) in the current frame with its surrounding area in the previous frame and attempts to find a match. The matching area is moved into the current macroblock position by the motion compensator module. The motion compensated macroblock is subtracted from the current macroblock. If the motion estimation and compensation process is efficient, the remaining "residual" macroblock should contain only a small amount of information. (4I2I, 2000)

d) Discrete Cosine Transform (DCT)

The DCT transforms a block of pixel values (or residual values) into a set of "spatial frequency" coefficients. This is analogous to transforming a time domain signal into a frequency domain signal using a Fast Fourier Transform. The DCT operates on a 2-dimensional block of pixels (rather than on a 1-dimensional signal) and is particularly good at "compacting" the energy in the block of values into a small number of coefficients. This means that only a few DCT coefficients are required to recreate a recognizable copy of the original block of pixels. (4I2I, 2000)

e) Quantization

For a typical block of pixels, most of the coefficients produced by the DCT are close to zero. The quantizer module reduces the precision of each coefficient so that the near-zero coefficients are set to zero and only a few significant non-zero coefficients are left. This is done in practice by dividing each coefficient by an integer scale factor and truncating the result. It is important to realize that the quantizer "throws away" information. (4I2I, 2000)

f) Entropy Encoding

An entropy encoder (such as a Huffman encoder) replaces frequently occurring values with short binary codes and replaces infrequently occurring values with longer binary codes. The entropy encoding in H.263 is based on this technique and is used to compress the quantized DCT coefficients. The result is a sequence of variable-length binary codes. These codes are combined with synchronization and control information (such as the motion "vectors" required to reconstruct the motion-compensated reference frame) to form the encoded H.263 bit stream. (4I2I, 2000)

g) Frame Store

The current frame must be stored so that it can be used as a reference when the next frame is encoded. Instead of simply copying the current frame into a store, the quantized coefficients are re-scaled, inverse transformed using an Inverse Discrete Cosine Transform and added to the motion-compensated reference block to create a

reconstructed frame that is placed in a store (the frame store). This ensures that the contents of the frame store in the encoder are identical to the contents of the frame store in the decoder (see below). When the next frame is encoded, the motion estimator uses the contents of this frame store to determine the best matching area for motion compensation. (4I2I, 2000)

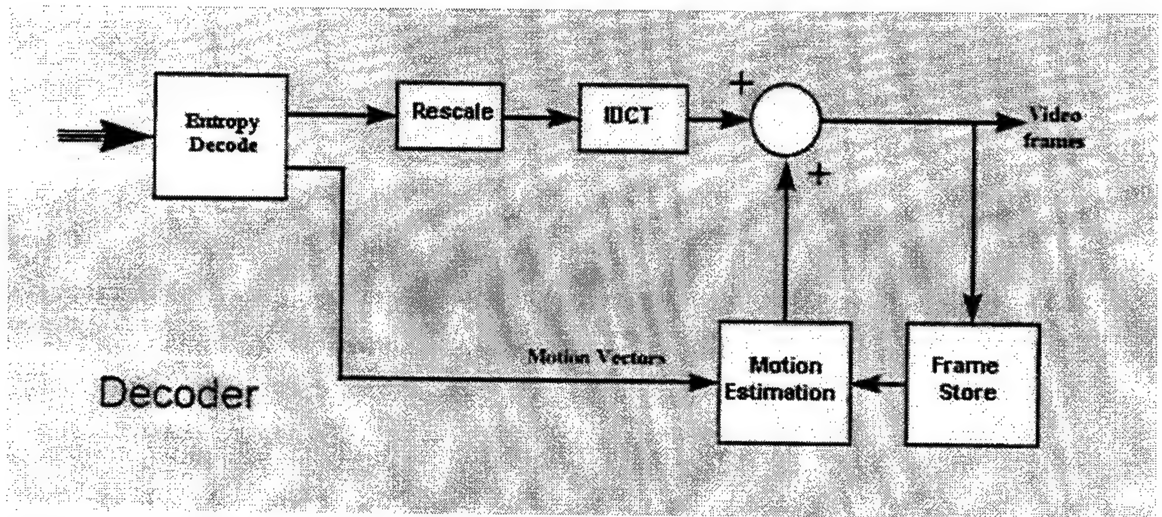


Figure 2.3 H.263 Decoder

h) Entropy Decode

The variable-length codes that make up the H.263 bit stream are decoded in order to extract the coefficient values and motion vector information. (4I2I, 2000)

i) Rescale

This is the "reverse" of quantization: the coefficients are multiplied by the same scaling factor that was used in the quantizer. However, because the quantizer

discarded the fractional remainder, the rescaled coefficients are not identical to the original coefficients. (4I21, 2000)

j) Inverse Discrete Cosine Transform

The IDCT reverses the DCT operation to create a block of samples: these (typically) correspond to the difference values that were produced by the motion compensator in the encoder. (4I2I, 2000)

k) Motion Compensation

The difference values are added to a reconstructed area from the previous frame. The motion vector information is used to pick the correct area (the same reference area that was used in the encoder). The result is a reconstruction of the original frame: note that this will not be identical to the original because of the "lossy" quantization stage, i.e. the image quality will be poorer than the original. The reconstructed frame is placed in a frame store and it is used to motion-compensate the next received frame. (4I2I, 2000)

l) Implementation Issues

Real-time video communications. Many issues need to be addressed in order to develop a video encoder and decoder that can operate effectively in real time. These include:

- Bit rate control. Practical communications channels have a limit to the number of bits that they can transmit per second. In many cases the bit rate is fixed (constant bit rate or CBR, for example POTS, IDSN, etc.). The basic H.263 encoder generates a variable number of bits for each encoded frame. If the motion estimation/compensation process works well then there will be few remaining non-zero coefficients to encode. However, if the motion estimation does not work well (for example when the video scene contains complex motion), there will be many non-zero coefficients to encode and so the number of bits will increase. In order to "map" this varying bit rate to (say) a CBR channel, the encoder must carry out rate control. The encoder measures the output bit rate of the encoder. If it is too high, it increases the compression by increasing the quantizer scale factor: this leads to more compression (and a lower bit rate) but also gives poorer image quality at the decoder. If the bit rate drops, the encoder reduces the compression by decreasing the quantizer scale factor, leading to a higher bit rate and a better image quality at the decoder. (4I2I, 2000)
- Synchronization. The encoder and decoder must stay in synchronization, particularly if the video signal has accompanying audio. The H.263 bit stream contains a number of "headers" or

markers: these are special codes that indicate to a decoder the position of the current data within a frame and the "time code" of the current frame. If the decoder loses synchronization then it can "scan" forward for the next marker in order to resynchronize and resume decoding. It should be noted that even a brief loss of synchronization can cause severe disruption in the quality of the decoded image and so special care must be taken when designing a video coding system to operate in a "noisy" transmission environment. (4I2I, 2000)

- Audio and multiplexing. The H.263 standard describes only video coding. In many practical applications, audio data must also be compressed, transmitted and synchronized with the video signal. Synchronization, multiplexing and protocol issues are covered by "umbrella" standards such as H.320 (ISDN-based videoconferencing), H.324 (POTS-based video telephony) and H.323 (LAN or IP-based videoconferencing). H.263 (or its predecessor, H.261) provide the video coding part of these standards groups. Audio coding is supported by a range of standards and will not be discussed here. Other, related standards cover functions such as multiplexing (e.g. H.223) and signaling (e.g. H.245). (4I2I, 2000)

- Software implementations. Functions such as motion estimation, variable length encoding/decoding and the DCT require a significant amount of processing power to implement. However, with recent developments in processor technology, it is possible to encode and decode H.263 video in real time on general-purpose processors such as the Pentium family. A software implementation must be highly optimized to achieve "reasonable" video quality (e.g. more than 10 frames per second, 352x288 pixels in each frame). This involves a number of steps such as choosing fast algorithms for processor-intensive functions, minimizing the number of move or copy operations and unrolling loops. In some cases assembly code routines (for example making use of Intel's MMX extensions) will further speed up operation.
- Hardware implementations. For high quality video, or in applications where a powerful processor is not available, a hardware implementation is the solution. A typical hardware CODEC might use dedicated logic for the computationally intensive parts of the system (such as the motion estimator/compensator, DCT, quantizer and entropy encoder) with a control module that schedules events and keeps track of the encoding and decoding parameters. A programmable controller is advantageous because many of the encoding parameters (such as

the rate control algorithm) can be modified or adapted to suit different environments. (4I2I, 2000)

C. MPEG

1. Background of MPEG

MPEG stands for the Motion Picture Experts Group. This is not affiliated with the motion picture industry; rather it is a group of computer scientists trying to make a standard for digital representation of video.

There are several MPEG standards and they are evolving constantly:

- MPEG-1: MPEG-1 was the original MPEG standard, designed exclusively for computer use. It allows for 320x240, 30 frames per second video.
- MPEG-2: MPEG-2 is a higher resolution version of MPEG-1 designed for digital television broadcast.
- MPEG-3: was designed for HDTV. However, HDTV is just normal TV with a higher resolution and frame rate, so this standard was folded into MPEG-2 and is no longer used.
- MPEG-4: MPEG-4 is a new standard for digital video that was approved in November of 1999. It is designed for use over low-bit-rate wireless and mobile communication systems. DCT cannot provide the required compression to operate over this type of network, so MPEG-4 does not force an encoding method. Instead, it leaves the choice of encoding method up to the designer.

- MPEG-7: MPEG-7 is yet another standard for digital video that has barely begun. The focus of MPEG-7 is supposed to be designing a representation for digital video that allows it to be stored and queried by content in a video database. [Rao 1996]

2. Why MPEG Over JPEG for Video?

JPEG, which stands for Joint Photographic Experts Group, is the standard for transmitting still images over digital networks. The JPEG algorithm is designed specifically for digital images. While MPEG does utilize JPEG to some extent, motion video has some additional properties that JPEG does not consider.

- Use and synchronization of multiple media streams, such as Video, Audio, and Closed-Captioning.
- Time relationship between frames.

Because video is displayed at 30 frames per second, even JPEG cannot give us the compression necessary to make digital video feasible. However, if we can exploit the relationship between successive frames (there will likely be little or no change between frames), we can compress even more. MPEG accomplishes this through Inter-frame Coding, Frame Types, Motion Estimation, Decoding versus Presentation Order, Independent versus Dependent GOP's, Bandwidth, Motion Estimation and Sub-sampling, and Error Handling. Each of the standards uses these techniques and the differences between them will be covered as each is discussed in more depth.

3. Inter-frame Coding

With 30 frames per second, you will naturally expect differences between successive frames of a video sequence to be relatively small. MPEG achieves a great deal of compression by exploiting the relationship between successive frames. Rather than encoding one initial frame and then sending only differences for all the remaining frames, MPEG uses a windowing approach. Windowing breaks up the video sequence into smaller subsequences and encodes differences only within a window, not between them. This is done for two reasons:

1. Protection from errors: What if you lose a frame in transmission? It is possible that the rest of the entire sequence could be useless without the windowing.
2. Random access and editing: How could you edit a compressed video sequence without having to decompress then re-encode it without windowing.

Each of these windows in MPEG is called a Group of Pictures (GOP). A GOP can be any length you like. There is none specified in the standard and a video sequence can contain GOP's of various lengths.

4. Frame Types

- I frames are intracoded frames. They do not depend on any other frames, you can think of them as JPEG images.
- P frames are predicted frames. They depend on the previous P frame or I.
- B frames are bi-directional frames. They can depend on either the previous or next I or P frame.

Because I and P frames are used to predict other P and B frames they are referred to as reference frames. [Laplante 1996]

5. Motion Estimation

Motion estimation is perhaps one of the most import considerations when examining the application of digital video to the intelligence process. This is due to the high motion characteristics of the tactical environment. Motion Estimation in MPEG operates on macroblocks. A macroblock is a 16x16 pixel range in a frame. There are two primary types of motion estimation, forward and backward. Forward prediction predicts how a macroblock from the previous reference frame moves forward into the current frame. Backward prediction predicts how a macroblock from the next reference frame moves back into the current frame.

Motion estimations operate as follows: First, compare a macroblock of the current frame against all 16x16 regions of the frame you are predicting from. Then select a 16x16 region with the least mean-squared error from the current macroblock and encode a motion vector, which specifies the 16x16 region you are predicting from and the error values for each pixel in the macroblock. This is done only for the combined Y, U, and V values. Subsampling and separation of the Y, U and V bands comes later.

There are four types of macroblocks:

1) Forward Predicted: (P and B only) predict from a 16x16 region in the previous reference frame.

2) Backward Predicted: (B only) predict from a 16x16 region in the next reference frame.

3) Bi-directional Predicted: (B only) predict from the average of a 16x16 region in the previous reference frame and a 16x16 region in the next reference frame.

4) Intracoded: (I, P, or B) are not predicted, the actual pixel values are used for the macroblock.

It is important to remember that P and B frames can contain intracoded macroblocks as well as predicted macroblocks if there is no efficient way to predict the macroblock.

In MPEG, the coding process for P and B frames includes the motion estimator, which finds the best matching block in the available reference frames. P frames are always using forward prediction while B frames use the bi-directional prediction--also called motion-compensated interpolation. B frames can use forward or backward prediction, or interpolation. A block in the current frame (B frame) can be predicted by another block from the past reference frame (B= A → forward prediction), or from the future reference frame (B= C → backward prediction, or by the average of the of two blocks (B= (A+C)/2 →interpolation).

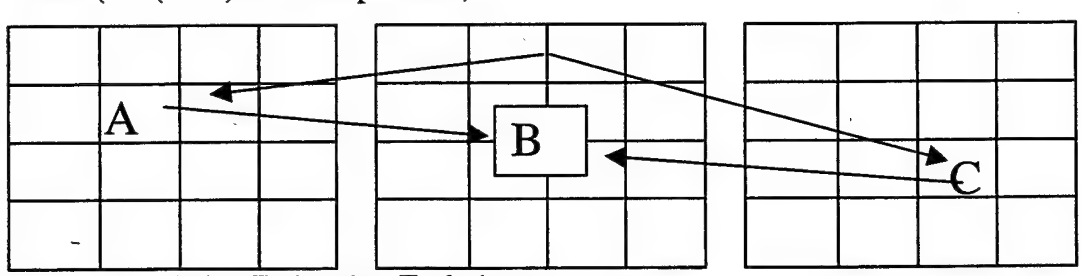


Figure 2.4 Motion Estimation Techniques

Motion Estimation is used to extract the motion information from the video sequence. For every 16x16 block of P and B frames, one or two motion vectors are calculated. One motion vector is calculated for P and forward- and backward- predicted B frames.

The MPEG standard does not specify the motion estimation technique; however, block matching techniques are likely to be used. Using a lock-matching motion estimation technique, the best motion vectors(s) are found, which specifies the space distance between the actual and the reference blocks. The difference between predicted and actual blocks, called error term, is then calculated and encoded using the DCT-based transform coding. The color image is first converted into YUV format. Each image consists of the luminance and two chromiance components. The luminaince has twice as many samples in the horizontal and vertical axes. [Rao 1996]

6. Decoding and Presentation Order

MPEG is actually used in decoding order rather than presentation order. Examples of both follow:

Presentation Order

I_1 B_2 B_3 B_4 P_5 B_6 B_7 B_8 P_9 B_{10} B_{11} B_{12} I_{13}

Decoding Order

I_1 P_5 B_2 B_3 B_4 P_9 B_6 B_7 B_8 I_{13} B_{10} B_{11} B_{12}

The reason for the difference is that in order to decode a predicted frame, all frames that it may be predicting from must be decoded first. Therefore, since $B_{2..4}$. This distinction becomes very important when you work with MPEG. (Ragahavan 1997)

7. Independent Versus Dependent GOP's

Independent GOP's do not depend on any frames of the previous GOP for prediction. Dependent GOP's depend on a reference frame from another GOP for prediction. Examples follow (in decoding order):

Case 1: GOP_2 (starts form I_{13}) is dependent on GOP_1

I_1 P_5 B_2 B_3 B_4 P_9 B_5 B_6 B_7 I_{13} B_{10} B_{11} B_{12}

Case 2: GOP_2 (starts from I_{13}) is not dependent on GOP_1

I_1 P_5 B_2 B_3 B_4 P_9 B_5 B_6 B_7 P_{12} B_{10} B_{11} I_{13}

To illustrate the difference, imagine trying to perform a simple edit operation that cuts out GOP_1, consequently removing P_9. If this happens, $B_{10...12}$ will not be able to be decoded since they depend on P_9. In the second case no frames in the second GOP depend on the first GOP, making this operation possible. As shown here, if you want to make a dependent GOP independent, end the first GOP with a P frame. (Raghavan 1997)

8. Bandwidth

Bandwidth is a major concern in digital video and the application of it in already over tasked command and control networks. Characteristics of MPEG that need to be considered for bandwidth management:

- I Frames require the most space, and give the least compression
- B frames require the least space and give the most compression

- P frames are in between

The following are Parameters of the MPEG Algorithms:

Format	MPEG	Video Parameters	Compressed Bit Rate
SIF	MPEG	352x240 @ 30Hz	1.2-3 Mb/s
EDTV	MPEG-2	960x486 @ 30Hz	7-15 Mb/s
HDTV	MPEG-2	1920x1080 @ 30 Hz	20-40Mb/s
Multimedia	MPEG-4	160x120 @ 30 Hz	9-64 Kb/s

Figure 2.5 MPEG Parameters

If an encoded stream is bigger than the available bandwidth, the encoder will quantizize more coarse (to increase compression) and re-encode the sequence. This is called feedback. The output of the encoder will be analyzed and re-encoded until it can fit the available bandwidth. This degrades quality of service through loss of resolution. [Laplante 1996]

D. STREAMING VIDEO TECHNOLOGY

MPEG-4 is an ISO/IEC standard developed by MPEG (Moving Pictures Experts Group), the committee that also developed MPEG-1 and MPEG-2. These standards made interactive video on CD-ROM and Digital Television possible. MPEG-4 is the result of another international effort involving hundreds of researchers and engineers from all over the world. MPEG-4, whose formal ISO/IEC is designation is ISO.IEC 14496, was finalized in October 1998 and became an International standard in 1999. (MPEG 1999)

MPEG-4 builds on the proven success of three fields:

- Digital Television
- Interactive graphics applications (synthetic content)
- Interactive Multimedia (World Wide Web, distribution of and access to content)

MPEG-4 provides the standardized technological elements enabling the integration of production, distribution and content access paradigms of the three fields.

1. Scope and features of the MPEG-4 Standard

The MPEG-4 standard provides a set of technologies to satisfy the needs of authors, service providers and end-users alike.

- For authors, MPEG-4 enables the production of content that has far greater reusability, has greater flexibility than is possible today with the individual technologies such as digital television, animated graphics, World Wide Web pages and their extensions. Also, it is now possible to better manage and protect content owner rights.
- For network service providers, MPEG-4 offers transparent information, which can be interpreted and translated into the appropriate native signaling messages of each network with the help of relevant standards bodies. The forgoing, however, excludes Quality of Service (QoS) considerations, for which MPEG-4 provides a generic QoS descriptor for the different MPEG-4 media. How this QoS is implemented is left up to the service provider.

- For the end user, MPEG-4 brings higher levels of interaction with content, within the limits set by the authors. It also brings multimedia to new networks, including those employing relatively low bit rate and mobile ones.

For all parties involved, MPEG-4 seeks to avoid a multitude of proprietary, non-interworking formats and players. (MPEG 1999)

2. Comparing and Choosing Streaming Video Technology

Now that the basics of how digital video is produced have been discussed the challenge of implementing a streaming video application is covered next. The streaming video industry has exploded over the past four years and three main players have come to the forefront. All of them claim to be the leader and each will be examined with the pluses and minus of each described. First one must define what is streaming video-- Is it all video on the Web or only video that is streamed through UDP (User Datagram Protocol, a protocol for the web that is different than HTTP). For this discussion we will define real streaming as UDP video and the usual HTTP version as progressive download. [Wagonner 2000]

The biggest difference is that true streaming only works when the bandwidth is large enough to play the video in real-time. Progressive download transfers at the available bandwidth and caches as much as is needed to your hard drive to act as a buffer before beginning playback. Progressive download usually ensures a higher-quality playback at any bandwidth, but with a potentially long delay. Complicating things is the hybrid of passing real time streaming video via HTTP if a firewall is unable to pass UDP data.

3. The Leaders

a) QuickTime

Apple's QuickTime, while the oldest, is a brazen newcomer to the group. It is the oldest digital video architecture around, serving as the foundation for the entire industry. It has been digital video, progressive download, for many years. Its support for true streaming only came out recently with QT v4.0.

Apple's QuickTime V4.1 was recently released and brought some advances with it. First, it includes support for SMIL, the same rich media that is the core of RealMedia. Second, it now supports streaming through an HTTP connection, which makes QuickTime as capable as RealMedia and Windows Media at getting through firewalls. Lastly, the Macintosh version has added Apple Script support to help with automated media creation. The native file format is a QuickTime file, .mov, or it can be .qt or .qti.

b) RealVideo

RealVideo, from RealNetworks, is another pioneering Web streaming format. RealAudio came out in 1994; RealVideo was added in 1997 with the 4.0 upgrade. Version 7.0 was recently released providing substantial upgrades from the G2 version of 1999. The greatest improvements come from better decoding performance, improved encoding technologies and a full player makeover. The native file format is RealMedia, or .rm or .ram

c) *Windows Media*

Formally known as NetShow, Windows Media is Microsoft's entry into the streaming Web video market. The newest to the fray for video, it is being pushed hard by Microsoft. Windows Media is a much simpler solution than QuickTime or Real because Microsoft doesn't position it as a complete solution, but rather the streaming audio and video component of a Web browser. However, it does what it does quite well. Recent innovation has focused on codec improvements and implementation of pay-per-view and authentication features through the Windows Media Rights Manager. The native file format is Advanced Streaming Format, or .asf

4. Video Codecs

Video Codecs are probably the single most important factor in determining what makes a great video technology. Bandwidth is still quite limited and trying to get high quality video to the desktop is like squeezing an elephant through a swizzle stick. To go from uncompressed digital video to 28.8 Kbps modem bit rate requires around a 12000:1 compression ratio. The bang for the bit of a codec is obviously critical to the user viewing the experience. Universal broadband will ease this situation someday, but in the near future we need the best performance from codecs as possible. [Wagonner 2000]

a) *QuickTime*

QuickTime utilizes several dozen codecs. The Sorenson Video codec is best suited for the web. It is flexible providing competitive quality over a wide range of data rates. The Basic QuickTime comes with a stripped down version of Sorenson, the

full version, suitable for professional-quality video, is available with the QuickTime Developer edition.

The Developer Edition of Quick Time has better quality overall and has features that can be used to tweak the video to maximize the video for publication. The most important feature is the ability to codec for progressive downloads through the use of VBR, variable bit rate support. This allows you to get a higher average quality with the smallest file possible.

Sorenson Developer v2.0 encodes four times faster than the basic and the latest version v2.1 has some speed enhancements for both the Apple and Intel MMX platforms, 100% and 33% respectfully. Even though the software has speed enhancements it still slower than the Real and Windows Media codecs. Sorenson's codec enjoys a deeper level of compression knowledge due to its many codec options. QuickTime also has the H.263 codec, a standard video conferencing codec, that can yield better results then Sorenson for high motion content at lower modem rates.

b) Real

Real utilizes only one modern video codec, Real G2 video codec. The G2 codec is based on video conferencing technology from Intel, providing high quality and fast encoding. Initially designed for the 28.8kbps data rate and Pentium MMX technology, it did not scale to broadband very well. To address this they have taken steps to assure that processors and bandwidth across the board enjoy the ability to view the video.

With Real's Scaleable Video Technology (SVT), slower machines do not have to decode all of the original image data, resulting in poorer quality but a smooth playback. On a powerful enough platform, any edge artifacts are filtered away and a smoother appearance results. Lastly, the G2 codec can interpolate, forward and backward prediction, between two frames, giving you the ability, with the application of enough processing power, to play the video back at a higher frame than which it was recorded.

Real G2 is not a WYSIWYG (What You See Is What You Get) codec, so those users on slower speeds may not be able to get the great video you would have on your high end rendering system. Always test your applications on the minimum platform you plan on supporting.

Real v.7.0 has improved on the G2 performance. First it has sped up the decoder, enabling the mid-range machines to get the full benefit of RealPlayer. The new encoder, currently in beta, will support a technique similar to VBR encoding. It examines the entire video stream and budgets its bit allotment to those frames with high levels of motion. You can also increase the buffer size, allowing the encoder more time to find the optimal bit allocation, but this will result in a delay in the clip starting. This is a worthwhile investment in time for the higher quality it yields.

c) *Windows Media*

The most important codec in Windows Media is the proprietary MPEG-4 v3. It is a great codec providing high performance and video quality over a wide range of data rates. It is a fast compressor but does not offer Variable Data Rate encoding. The

Windows Media Advanced Streaming File (.asf) is not the same as a MPEG-4 standard file, which is based on the QuickTime file format.

5. Multiple Data Rate Support

Each of the systems has the ability to link users with multiple data rates, allowing your video to play without the user having to specify the bandwidth.

a) QuickTime

QuickTime's approach to multiple data rate support is quite radical and time consuming. Instead of bundling multiple data rates in a single file, you create different files for each. This complicates encoding and does not address the problem of fluctuating bandwidths. The upside is you can provide different content for different users based on platform and bandwidth.

b) Real

Real's SureStream technology lets you put multiple tracks in a single file. You can vary every parameter for any given bandwidth except resolution, video and audio codec, frame rate, and data rate. This allows you to bundle a modem and a broadband stream together. It also supports the bundling of older Real version streams within the SureStream, allowing those with older viewers to see something.

c) Windows Media

Windows Media has a limited way of handling multiple data rates called Intelligent Streaming. Multiple video tracks are encoded in a single file, with only the data rate parameter changing. You can vary the codec or frame rate. The down side is

there is only one audio track and this makes Intelligent Streaming not true multiple data rate support. However, this is a useful tool for handling network fluctuations by providing a backup stream. The user will experience the lower quality of video only for the time the connection has been degraded. [Microsoft 2000]

As has been highlighted through the discussion of the various standards and techniques for producing source video and the many choices for converting that video into a format suitable for streaming, the streaming of high quality video, even under optimal conditions is not easy. Add to this equation multiple bandwidths and high motion, the challenge of high quality video to the user becomes a daunting task. These factors give an appreciation for the unique challenges of implementing these technologies in a tactical environment.

E. HOW HUMANS SPATIALLY PERCIEVE

After examining the technological aspect of generating a video stream logically it would be appropriate to examine briefly the impact that "how" one views or experiences something affects the way one perceives the experience.

1. Active Versus Passive Viewing

In a study of spatial perception, conducted by Patrick Pe'ruch, Jean-Louis Vercher and Gabriel M.Gauthier, of a subject's ability to learn a graphically displayed wall limited environment they determined that performance was better for active exploration than for passive exploration. A direct link was drawn between the level of performance and the level of spatial knowledge, and confirmed the importance of active motor behavior

combined with active perception to extract invariants from the environment.(Per'uch, 1995)

An observer moving in an unknown environment aquires spatial knowledge of the environment, which is progressively improved as the exploration duration and/or the number of displacements increase. When an observer moves through a real space, such as driving, information on self-generated displacement is available from different sensory receptors, these senses are diminished in streaming video. The importance of these sensory modalities has been documented with vision being the most dominant. This visual flow is critical to characterizing the observer's displacement through the environment. The nature of the displacement (active/passive) and the type of visual information (continuous sweeping/successive fixed frames) may also result in significant differences in acquisition of spatial knowledge. The more active and more continuous the viewing the better the acquisition of spatial knowledge. (Per'uch, 1995) It is for these reasons it is proposed that quality of service of streaming video (data rate/frames per second) ultimately determines the usefulness of implementing this technology to assist the commander.

THIS PAGE INTENTIONALLY LEFT BLANK

III. INFORMATION SYSTEMS TO SUPPORT STREAMING VIDEO

In order to implement the proposed application of streaming video to the commander in the field there needs to exist a capacity to provide high speed data transfer from the source to the commander. The infrastructure of hard wired systems exists to the commander, but the means of injecting this signal wirelessly into the existing network is tenuous at best. It is this reach back capability that is the lynch pin in successfully implementing the proposed applications. This chapter will examine those existing and near future wireless systems that may meet this need in the future. This examination does not attempt to be exhaustive but serves to highlight the need and performance of those systems discussed.

A. CURRENT WIRELESS INFORMATION SYSTEMS

There are a number of systems available to transmit information around the battlespace, few of which have the needed bandwidth at this time to stream video. This section will be a sampling of systems that exist now and ones that are under development that are capable of carrying a signal at sufficient bandwidth to allow streaming of video. It will give an overview of the system but will not address the specifics of how the video will be injected into the network.

1. AN/PSC-5 (V) Shadowfire

The Shadowfire radio is currently the only man portable radio that can be fielded in the mission of streaming video back to the commander as described in this thesis. But even this is not without some considerations that will be discussed in this section.

The Shadowfire radio capitilzes on the AN/PSC-5 Spitfire's expandable modular architecture to satisfy user's requirements for full AM, FM, and FSK communications in the 30-512 MHz frequency range. It has high-data rate options of 76.8 Kbps Line of Sight and 56Kbps SATCOM.

One consideration that was highlighted in a conversation with the systems engineer from Raytheon, the manufactuer, is that the Carrier to Noise ratio to attain the higher data rates needs to be close to the link. The current MIL-STD-188-181B requires a 1E-5 Bit Error Rate and at 56Kbps is 61dB-Hz. This requires an amplifier or large antenna. The problem is also excacerbated by low power transponders and low elevation angles in various theatres. To quote Mark Reese, RF Systems Engineer, Raytheon Corporation " This quickly moves this out of the man-portable arena into the vehicular transported world."

B. NEAR FUTURE WIRELESS SYSTEMS

The need for broadband communication channels has not been lost on the government or commercial sectors. In response to this there has been a prolifieration of satellite based systems being developed. Two systems, one commerecial and one government sponsored, are discussed as solutions to the problem of limited bandwidth.

1. Military

a) MILSTAR II

Milstar II is the next generation military satellite communication system, designed to serve the National Command Authority and the Unified and Specified

commanders and their operational forces. Milstar II will be the Department of Defense's core command and control communications system for U.S. strategic and tactical combatant forces in hostile environments well into the next century.

Milstar II will provide a combination of capabilities unmatched by any other satellite communication system. These capabilities include worldwide, secure, survivable, highly jam resistant communications; satellite-to-satellite communication; autonomous operation; the ability to reposition to meet theater requirements; and the ability to provide direct support to mobile forces. These capabilities are achieved through first-time use of extremely high frequency (EHF) and advanced processing techniques.

The Milstar II payloads perform extensive on-board processing of the uplink and downlink waveforms for efficient on-orbit resource use and maximum antijam performance. On-board signal processing ensures full interoperability among the military services and other users who operate terminals on land, sea, and air.

Often described as a switchboard in the sky, the Milstar II payloads have on-board computers that perform communications resource control. Milstar II responds directly to service requests from user terminals Without satellite operator intervention, providing point-to-point communications and network services on a priority basis.

EHF provides natural jam resistance, a function that is further enhanced by processing techniques on board the spacecraft which allow communications to be independent of ground relay stations and ground distribution networks. Automatic management of the satellite communication network will allow services to be established in minutes, instead of the hours and days needed by current systems. EHF also allows use

of smaller and more mobile terminals that will be installed on aircraft, ships, and land vehicles. Man-portable systems are also being developed.

DoD recommended and Congress concurred that a Medium Data Rate (MDR) payload should be added to the Milstar satellite to support tactical users with an increase in communications capacity. The MDR payload will be added to the third and all subsequent satellites. Contract award for development of the first Milstar II LDR/MDR satellite was in October 1992 following the Defense Acquisition Board program review. It is this MDR capability that will allow the MILSTAR II satellites to handle 4.8kbps to 1.544Mbps throughput. The development of man-portable terminals is still under development.

2. Commercial

a) Teledesic

Teledesic is building a global, broadband Internet-in-the-Sky™ network. Using advanced satellite technology, Teledesic and its partners are creating the world's first network to provide affordable, worldwide, "fiber-like" access to telecommunications services such as computer networking, broadband Internet access, interactive multimedia and high-quality voice. On Day One of service, Teledesic will enable broadband connectivity for businesses, schools and individuals everywhere on the planet. The Teledesic Network will accelerate the spread of knowledge throughout the world and facilitate improvements in education, health care and other crucial global issues. (Teledesic, 2000)

- Network Capacity/Access Speeds.The Teledesic Network is designed to support millions of simultaneous users. Multiple manufacturers will offer a family of user equipment to access the network. Using "standard" user equipment, most users will have two-way connections that provide up to 64 Mbps on the downlink and up to 2 Mbps on the uplink. Higher-speed terminals will offer 64 Mbps or greater of two-way capacity. Sixty–four Mbps represents access speeds more than 2,000 times faster than today's standard analog modems. (Teledesic, 2000)
- User Equipment. The Teledesic Network's low orbit eliminates the long signal delays normally experienced in satellite communications and enables the use of small, low-power user equipment to send and receive data. The fixed user equipment will mount on a rooftop and connect inside to a computer network or PC. Mobile applications are still being developed. (Teledesic, 2000)

Teledesic terminals communicate directly with the satellite network and support a wide range of data rates. The terminals also interface with a wide range of standard network protocols, including IP, ISDN, ATM and others. Although optimized for service to fixed-site terminals, the Teledesic Network is able to serve transportable and mobile terminals, such as those for maritime and aviation applications. (Teledesic, 2000)

Most users will have two-way connections that provide up to 64 Mbps on the downlink and up to 2 Mbps on the uplink. Broadband terminals will offer 64 Mbps of two-way capacity. This represents access speeds up to 2,000 times faster than today's standard analog modems. (Teledesic, 2000)

The ability to handle multiple channel rates, protocols and service priorities provides the flexibility to support a wide range of applications including the Internet, corporate intranets, multimedia communication, LAN interconnect, wireless backhaul, etc. In fact, flexibility is a critical network feature, since many of the applications and protocols Teledesic will serve in the future have not yet been conceived. (Teledesic,2000)

IV. DETERMINING EFFECTS OF QUALITY OF SERVICE ON SPATIAL PERCEPTION

In an effort to answer the primary question posed by this thesis an evaluation of the effects of quality of service on spatial perception was conducted. The need for information to the commander is a focal point of Joint Vision 2020 and every effort is being made to leverage technology to improve the commander's OODA loop. Video from the forward deployed forces to the commander is proposed to shorten this cycle. The technology to transmit this video is advancing and is allowing for video to be transmitted at lower and lower bit rates. How low is low enough but not too low? The bottleneck for the video technology is the wireless transmission technology. This bottleneck and the impact it has on the usefulness of the video is central to this thesis.

A. METHODOLOGY

The evaluation consisted of four groups of subjects watching a video stream at a quality of service level consistent with present and near future data rates for information systems. While watching the video stream the subject was asked to plot their location and orientation within the environment at specific time intervals on a floor plan of the environment. After watching the video stream the subject was asked to place various objects from the environment on the same floor plan. The subject was then asked to repeat the tasks to determine if there was a learning effect. This evaluation was consistent with the proposed implemention of these technologies for the commander in the field.

1. Video Content

The video that the subjects viewed was from the perspective of weapon-mounted camera on an individual as they enter and proceed to tactically search a building. The building was one that the subjects had no familiarity and only limited experience with the floor plan. The video was one minute and thirty seconds long.

2. Video Streams

The video streams were of a quality of service that is comparable to the expected data rates for existing and near future systems. The data rates simulated are 1.5Mbps, 256Kbps, 78Kbps, and 20Kbps. The streams were created by the using the Windows Media Encoder. Each stream was encoded for optimal transmission at the bit rate that it is intended to simulate. All video streams were encoded identically except for the targeted data rate. The below screen captures in Figures 5.1 and 5.2 are from the *Windows Media Encoder* and are representative of the settings used for this evaluation.

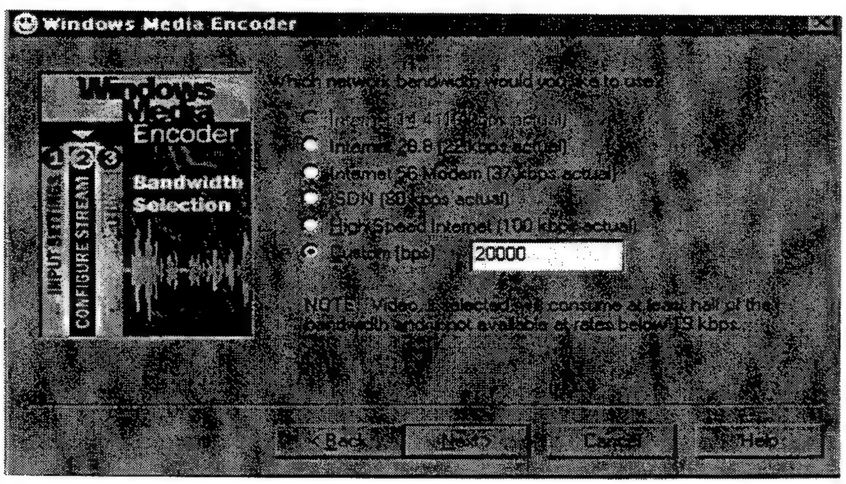


Figure 5.1 Windows Media Encoder

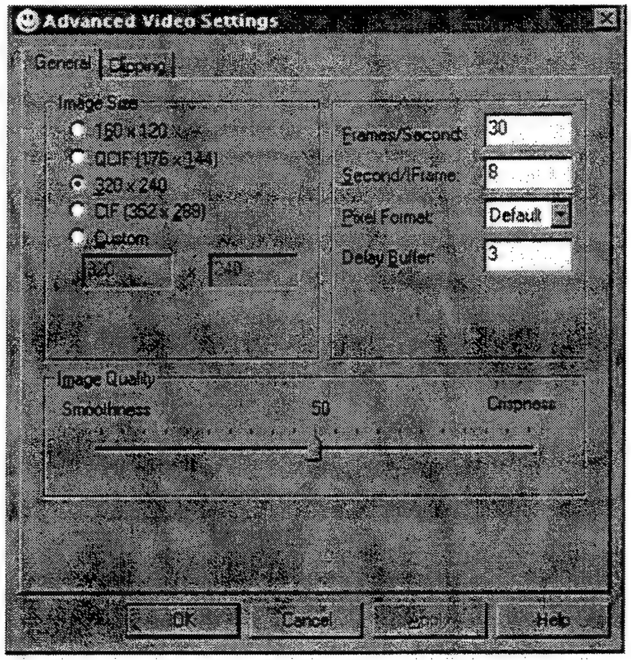


Figure 5.2 Advanced Video Settings

The resultant video streams were stored on the hard drive of the PC used to view them and had the characteristics depicted in Figure 5.3. One interesting characteristic that could not be explained after repeated encodings was that the highest bit rate video actually was encoded at a lower frame rate then the next lowest rate.

Video Stream	Data Rate	File Size	Encoded Frame Rate	Avg. Actual Frame Rate
T-1	1.5Mbps	17.24MB	30 fps	17.9 fps
VTC	256Kbps	2.78MB	30 fps	21.37 fps
Shadowfire	78Kbps	869KB	30 fps	6.71 fps
Minimum rate	20Kbps	230KB	30 fps	1.43 fps

Figure 5.3 Video Stream Characteristics

3. Viewing Method

The subjects viewed the selected video stream using *Windows Media Player v. 6.4* on a PC. The PC was an Intel Pentium Based system running at 398Mhz with 128MB

RAM and a 19" Viewsonic Monitor. The subject was seated in front of the PC and was provided a floor plan of the building (Appendix A). The floor plan was mounted on a small board and with the starting point of the video indicated on it. Each video is viewed at the same size of 320x240 pixels.

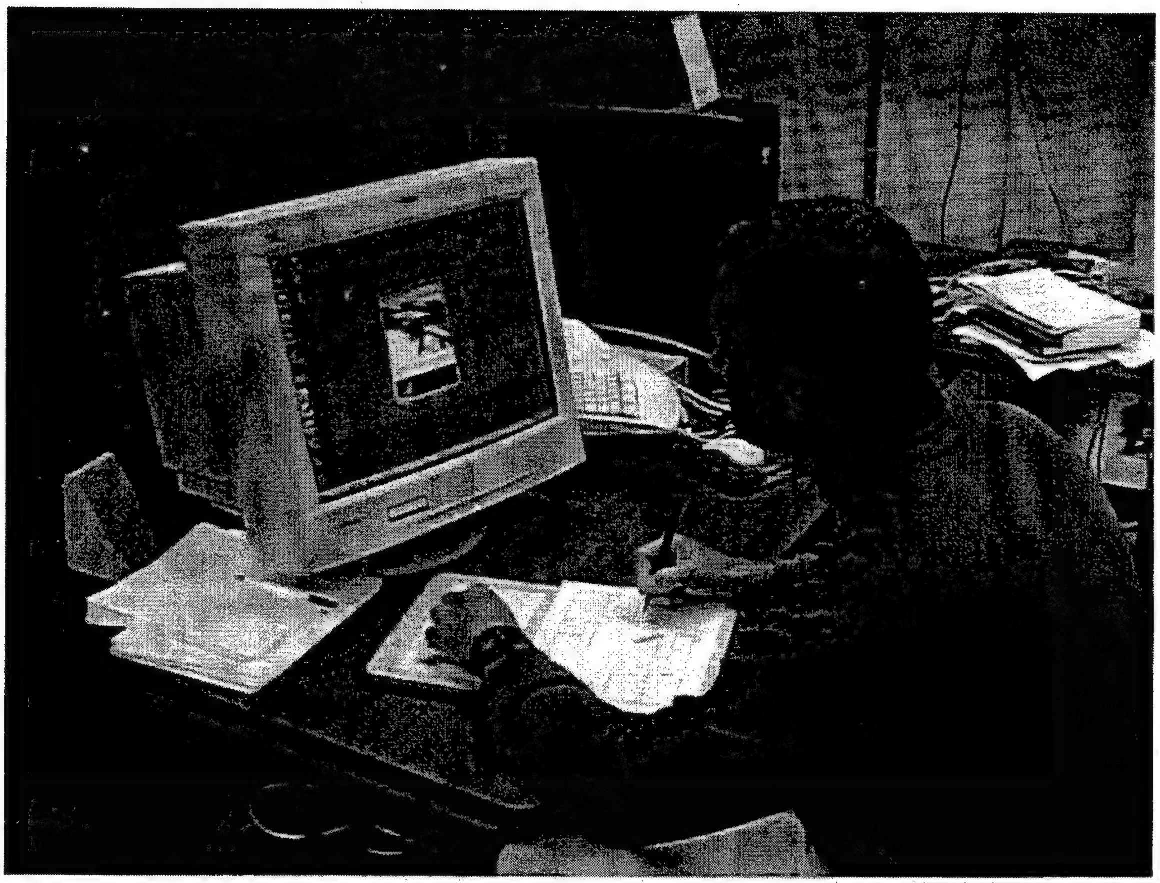


Figure 5.4 Subject Undergoing Spatial Perception Task

4. Objects from the Environment

Before the video stream was viewed the subject was provided frame captures from video that have a unique object depicted in them (Appendix B). There were five pictures. The subject was allowed to look over the objects for as long as they felt necessary. After the immediate end of the video stream the subject was asked to place the objects on the

floor plan where they felt they were located. They did this by placing a post-it arrow indicating this location on the floor plan.

5. Instructions for the Subject

The subject was read a scripted set of instructions (Appendix C) explaining the details and purpose of the experiment. Each of the two tasks, Spatial Orientation and Object Location, was explained and then the subject was asked if they had any questions. Once questions concerning the conduct of the experiment were answered the subject was put through their tasks.

6. Post Experiment Survey

At the conclusion of the second attempt at the tasks the subject was asked a series of demographic and subjective questions concerning the tasks. They were asked:

a) Branch of Service?

b) Years of Service?

c) Did they find the task of maintain their spatial perception hard?

d) On a scale of 1-6 with 6 hardest, how hard?

e) What could have been done to make their task easier?

7. Assumptions

In order to facilitate the determination of best-case minimum bandwidth requirements some assumptions had to be made. The first was that there would be no degradation of the data rate during the viewing of the video in the field. Secondly, to optimize the quality of the stream it was encoded for the target bit rate, removing any

frames that could have slowed the transmission. Lastly, that each subject would do their best while completing the tasks.

B. SAMPLE GROUP

1. Source

The sample population consists of students and staff from the Naval Post-Graduate School. They were a sample of convienience, selected on the basis of who would be willing to spend the twenty minutes it took to participate in the experiment.

2. Years of Service

In order to gauge the experience level of the sample the years of service was determined. The average years of service was 9.9 years with a standard deviation of 5.5. This standard deviation is large but in the case of this sample it indicates that the sample had a good distribution and is reflective of the levels of the larger population the sample came from, NPS students.

3. Difficulty of Task

In order to gauge the perceived difficulty of the task the sample was asked to rank the tasks from 1 to 6 with 6 being impossible. After completing the task the subjects ranked the task as in Figure 5.4.

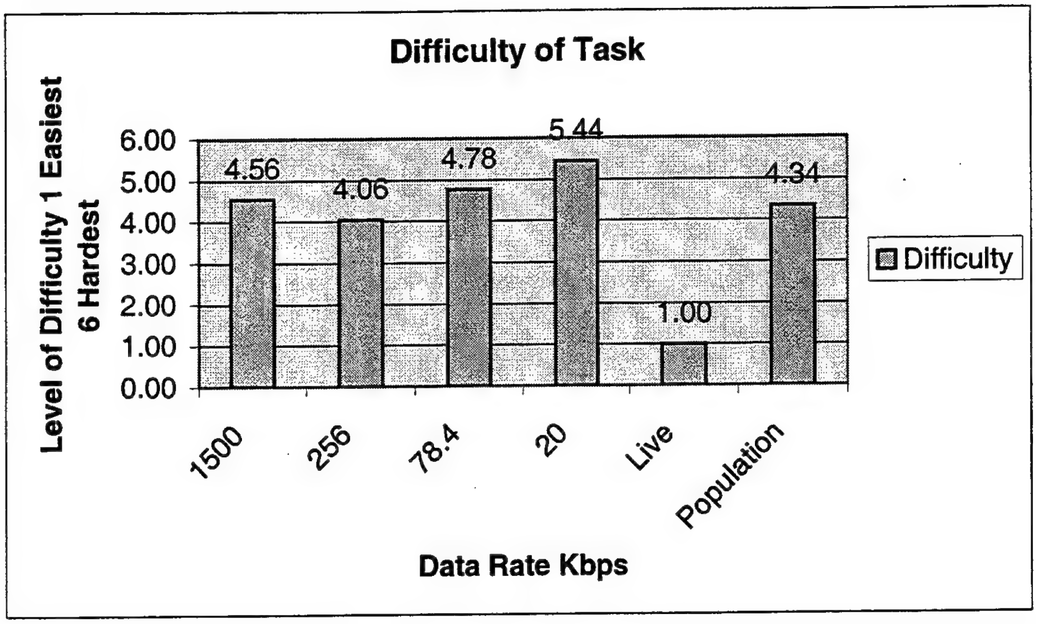


Figure 5.5 Task Difficulty

It is indicated from the lower perceived difficulty of the task at the highest frame rate that frame rate has some impact on the spatial perception task.

C. EVALUATION OF RESULTS

1. Spatial Perception

The subjects were asked to indicate their location and orientation within the environment while they watched the streaming video. The subjects were asked to indicate this with an arrow or mark. The orientation aspect of the task was not evaluated but was included to force the subject to be more exact in the placement of their location.

a) Determination of Results

The results were determined by evaluating the amount of linear distance of the subjects mark from a circle on an overlay of the floor plan that represented 4-foot diameter circle. This circle allowed for an error of four feet to be counted as zero. The

distance differential was measured in millimeters on the subjects floor plan by using a transparency of the floor plan with the actual locations indicated on it. This measuremnt is consistent and normalized with the floorplan. Each bit rate had a different overlay to eliminate any latency from encoding.

2. Objects Within the Environment

The subjects were asked to perform a secondary task of looking for a series of five objects within the environment and then placing them where they thought they were in the environment.

a) Determination of Results

The results for the placement of the objects from the environment was based on whether the object was seen and if it was placed in the correct room. If a subject placed an object in the environment but did not place it in the correct room it was not counted as having been correctly observed. Of the objects, two had multiple locations and credit was given for either location.

D. OBSERVATIONS

After the data was collected and the errors for each subject, run, and location were totaled, a determination was made to sum all the errors for each run to mitigate the compounding effect of an error in the spatial perception. When starting the examination of the data collected it was determined that there are four possible significant predictors for the total error. These predictors are the video(bandwidth), run, service, and years of service.

1. There is a Difference in Results Based on Civilian Versus Military Sample

By examining the box plot of the variances based on population it can be determined that there is a difference between the civilian population and the military population of the sample. Due to the small sample size of the civilian population this variance is not a significant factor in the results of the experiment. The Y-axis represents the total error for the spatial perception task and the X-axis is the service of the subject (U=United States Marine Corps, C=Civilian, N=Navy). The p-value indicates that these results are not from chance.

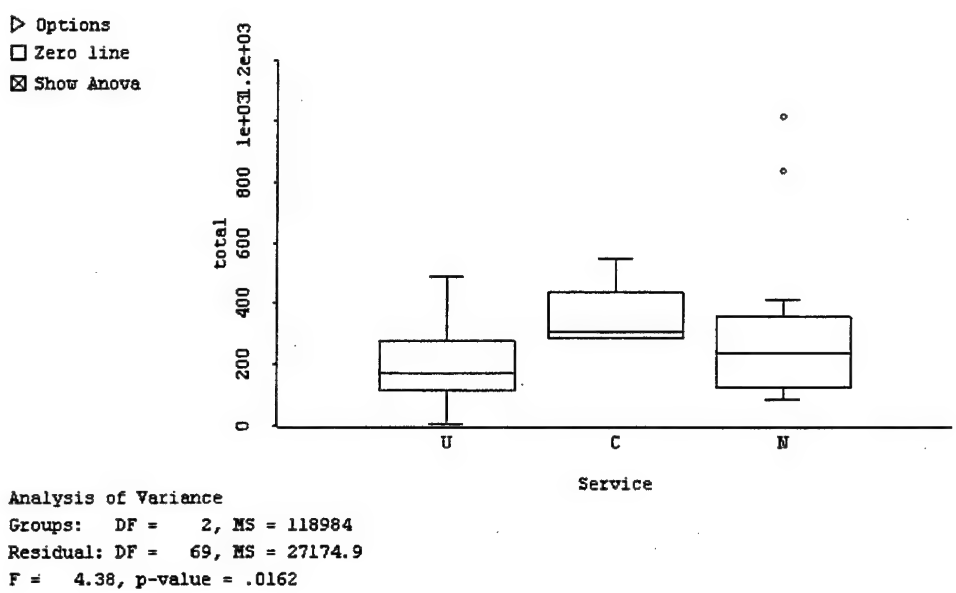


Figure 5.6 Analysis of Variance by Population

2. Experience Does not Impact the Results

Upon exmination of the experience of the sample in relation total error there is no statistical relevance to the amount of total error and the years of service in the sample

population Figure 5.6. TheY-axis is total error for the spatial perception task and the X-axis is Years Commissioned Service. The lines indicated on the chart are fitted lines based on the predicted values for the sample. The dotted line represents the predicted error for the Navy sample and shows a slight trend down in error as experience goes up. For the USMC sample, the predicted error represented by the solid line is flat, indicating no difference based on experience. Due to the small number of the Civilian in the sample no fitted line is shown.

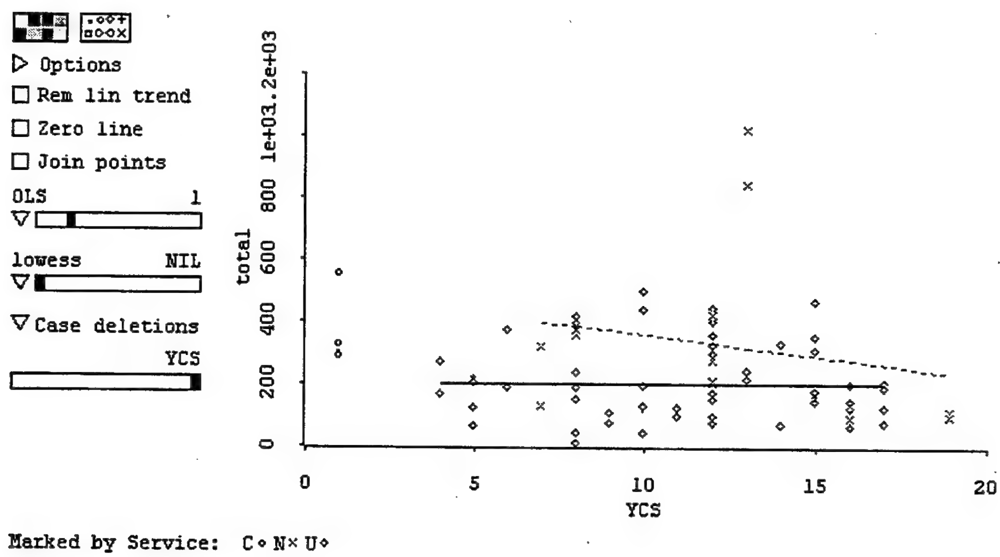
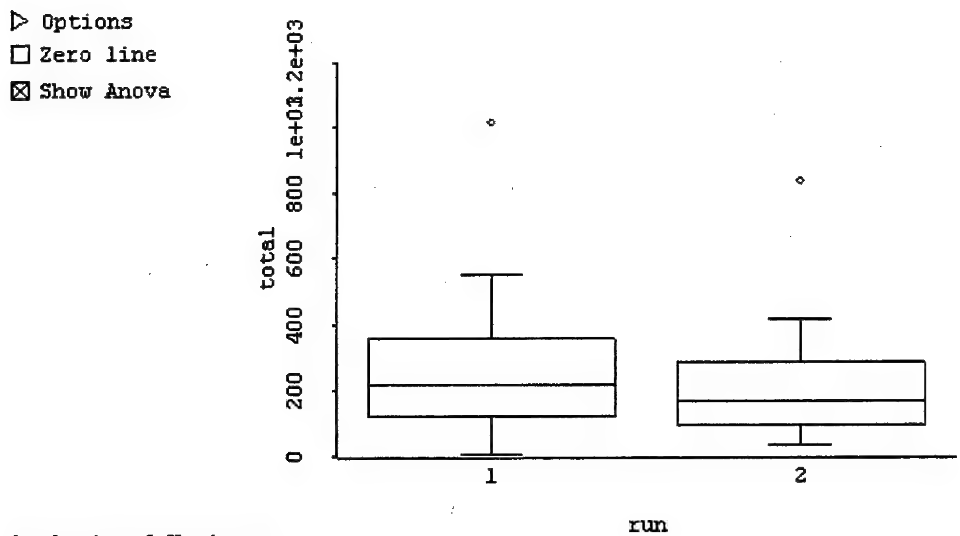


Figure 5.7 Total Error in Relation to Years Experience

3. There is No Learning Effect

Upon examination of the box plot, Figure 5.7, of the total error,Y-axis, and the runthat the subjects attempted, X-axis, there is no statistically significant difference which can be attributed to a learning effect. However, there is a trend of some improvement which is indicated by the compression of the box plot. The p-value of .13 indicates that there is a possibility of a learning effect from the repetition of viewing the same

environment but is inconclusive as to learning effect for different environments. The p-value indicates that this compression might also be from chance.




```
Analysis of Variance
Groups:   DF =    1, MS = 68511.7
Residual: DF =   70, MS = 29207.5
F =   2.35, p-value = .1301
```

Figure 5.8 Total Error in Relation to Run

4. Video Bandwidth Affects Total Error

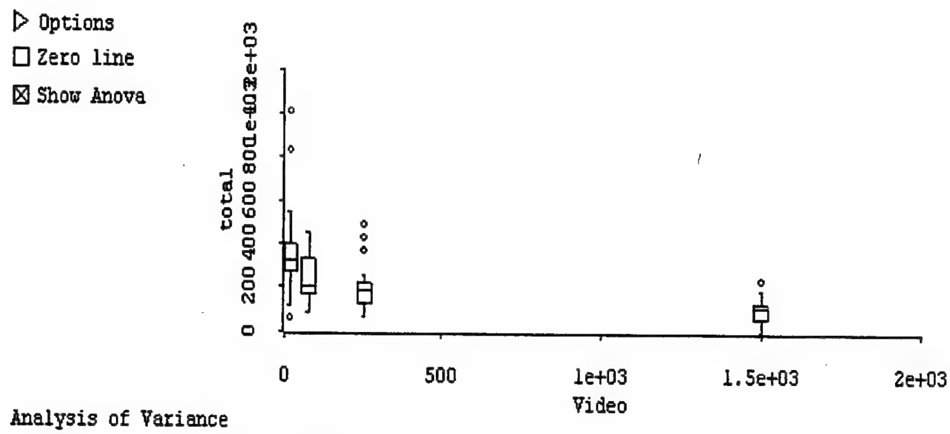
Through the running of a normal regression model in the Log scale it can be shown that the video bandwidth is a significant factor in the total error a subject had. Through a sequential analysis of variance it is indicated that the video is the most significant factor. The data listed in Figure 5.8 is the results of this analysis.

Total	Change				
Predictor	df	RSS	df	RSS	MS
Video	70	27.1773	1	14.5988	14.5988
run	69	26.1779	1	0.999456	0.999456
Subject	68	24.8859	1	1.29196	1.29196
YCS	67	24.7697	1	0.116221	0.116221
{F}Service	65	24.7516	2	0.0181105	0.00905524

Residual	65	24.7516	0.380794

Figure 5.9 Sequential Analysis of Variance

The Regression Sum of Squares of 14.5988 is an indicator that video is the most significant factor. All other factors have RSS's of less than two which indicates that they are not significant factors. Figure 5.9 represents the significance of bandwidth on error rates. The zero p-value indicates that there is no doubt that bandwidth affects the error rate.



Analysis of Variance
Groups: DF = 3, MS = 218697
Residual: DF = 68, MS = 21425.7
F = 10.21, p-value = .0000

Figure 5.10 Bandwidth Affect on Error

The p-value of the fit values vs. the residuals is not significant which indicates there is no curvature and allows for the development of a linear model to predict the total error for a given bandwidth. As is shown in Figure 5.10 the p-value is too large for us to accept the hypothesis that there is curvature in the model.

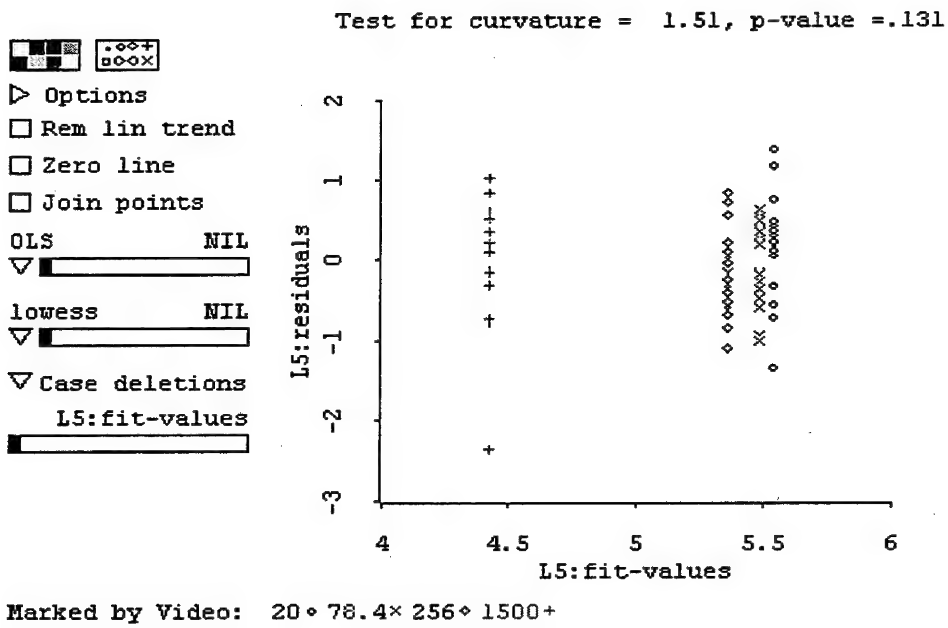


Figure 5.11 Residuals vs. Fitted Values in a Log Scale

The data listed in Figure 5.11 shows the values that can be used to predict the total error for a given bandwidth.

	Estimate	Std. Error	t-value
Constant	5.55819	0.0925330	60.067
Video	-0.000744717	0.000121447	-6.132

Figure 5.12 Date for Regression Model to Predict Total Error

Using this data a model for predicted error can be determined.

$$\text{Log(Total)} = 5.5819 + (.000744717)(\text{video kbps})$$

Using this formula one can determine the required bandwidth to give a commander the requested error rate.

$$\text{Log(Bandwidth)} = \frac{\text{Log(Total Error)} - 5.5819}{.000744717}$$

E. ANALYSIS OF SECONDARY TASK

The secondary task of locating objects within the environment was for most subjects impossible to do and still maintain their spatial perception. Figure 5.12 shows the burden that maintaining spatial perception detracts from other tasks. No subject was able to place more then 50% of the objects, except those that actually performed the tasks within the environment.

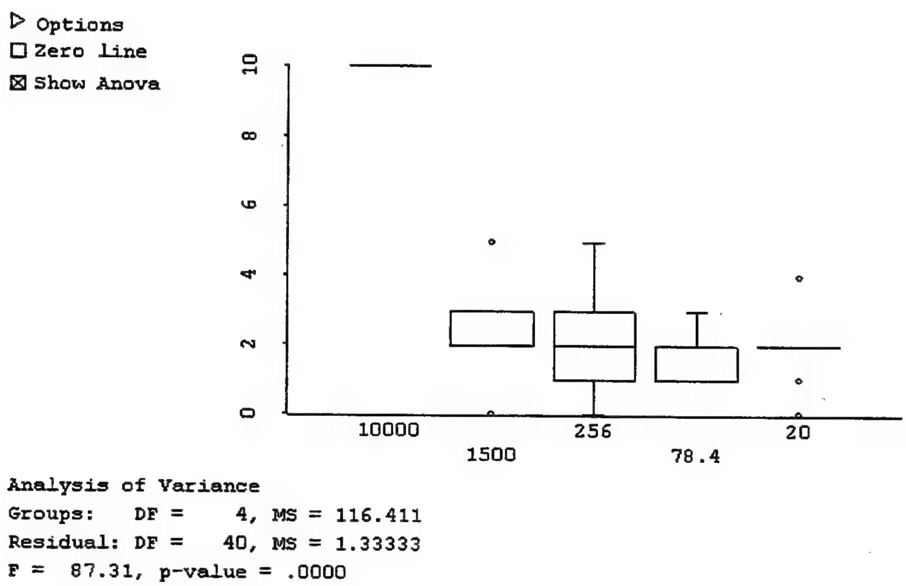


Figure 5.13 Object Recognition and Data Rate

F. SUBJECTIVE OBSERVATIONS OF WHAT IMPACTS RESULTS

1. Tracing the Route

When observing the subjects as they viewed the video it was observed that subjects who tried to actually trace their route through the building seemed to have more difficulty maintaining their spatial awareness. This, based on comments by the subjects, can be attributed to the fact that the video never stopped moving as they traced their

route. This lends valuable insight into the idea that the user must not have any distractions from the video if they are to be able maintain their spatial perception. This highlights the need to cache the video for further analysis and repetitive viewing in a less time compressed atmosphere.

2. Quick Marking

The technique of marking quickly and not making the mark "perfect" and focusing the video allowed some subject to keep their heads up and oriented while those who spent more then one or two seconds marking would get disoriented and would have trouble getting reoriented.

3. Pitch, Yaw, and Linear Movement

Between location 3 and location 4 as the camera moved out of an office space it panned down as it turned and moved laterally. This combination of pitch, yaw and linear movement confused each subject uniformally. Those with higher bandwidth managed to get reoriented faster while those at the lower bandwidths never caught up. This is best illustrtated by looking at the differential errors before and after location three. In figure 5.13 this is illustrated.

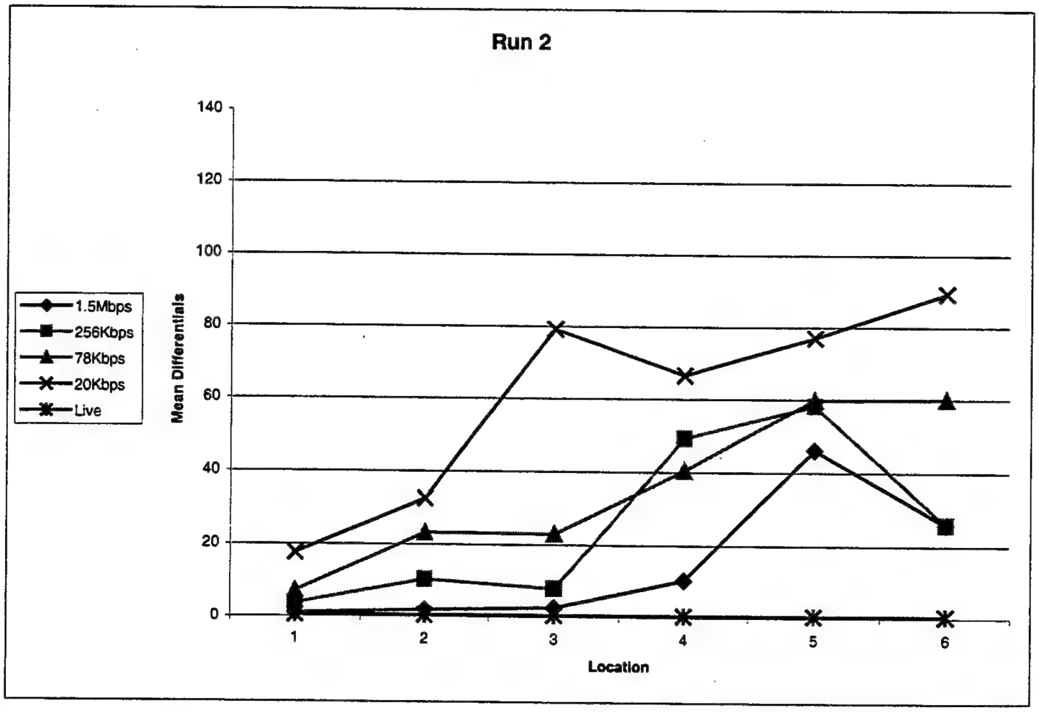


Figure 5.14 Differential by Location Run 2

G. RECOMMENDATIONS

This research brought to the forefront some significant isssues with regard to streaming video for the commander.

1. The bandwidth for streaming video, as indicated from the results of the experminent, has to be at a minimum of 256Kbps. More imporatantly , the resultant frame rate for any bandwidth needs to be at least 22 frames per second. Any lower and there is a severe degradation in the usefulness of the video.

2. The video stream needs to be cached for future analysis by those specifically trained in that skill. Without the caching and cataloging of metadata with the digital images the full potential of video will not be

realized. This is substantiated by the poor performance the subjects in the first run of the experiment.

3. The feasibility of mounting a camera on a weapon or helmet is very low and not recommended due to the high amount of pitch, yaw, and linear movement associated with an individual moving tactically through an environment. This problem compounds the effect of low frame rates on the usefulness of the video. A possible application of streaming video is the deployment of remote stationary video sensors to assist the commander with situational awareness.
4. Great consideration needs to be given to why and how video technologies are being fielded and to what level of command. There is a potential for a commander to "get lost in the weeds" and become overwhelmed to the point of degrading his situational awareness and decision making process from information overload.

H. FUTURE RESEARCH

Through the research conduct for this thesis there were many issues that were raised that could be the subject of future research.

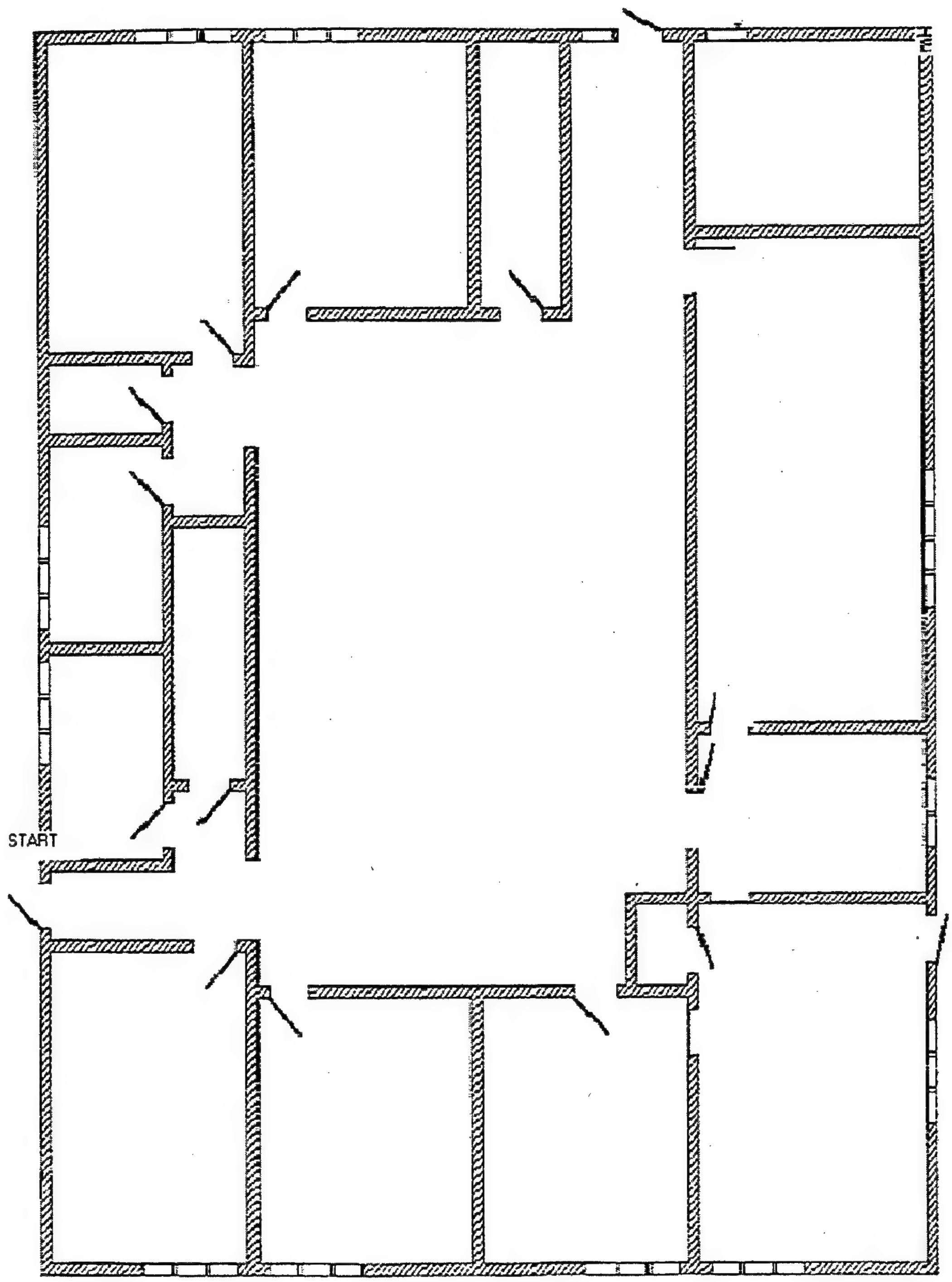
1. Benefits analysis of the implementation of a stereo audio feed to accompany any video to the commander and the impact this would have on the bandwidth overhead.

2. Development of an effective equipment suite to include the uplink to the wireless information system and field testing of the unit with a live feed from source to user.
3. Requirement generation and system development of information infrastructure required to manage multiple video streams from multiple platforms on the battlefield.
4. Development of proposed doctrine addressing the fielding and management of video technologies, from all sources, for the warfighter.

I. CONCLUSION

After examining the various video technologies available and developing a simulation of streaming video through the wireless information systems presently available and in the near future it is indicated that video bandwidth, which translates to the quality of service, is the most significant factor in determining the usefulness of a video stream. Without a high quality of service, the value to the commander in terms of shortening his decision making loop is minimal and may even degrade it. Streaming video may have a place on the battlefield; it just might not be on the commander's desktop.

APPENDIX A - FLOOR PLAN

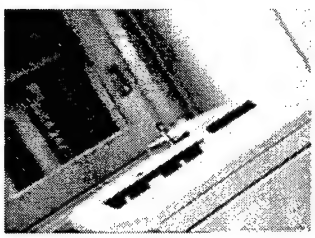


THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B - OBJECTS FROM ENVIRONMENT

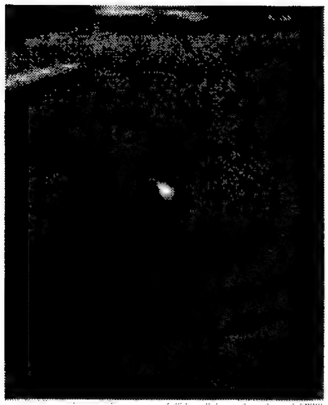
Floor Heater

Laser Printer



Fire Extinguisher

Floor Safe



Water Fountain

APPENDIX C - INSTRUCTIONS TO SUBJECTS

The video clip you are about to watch is a simulation of a video stream being transmitted back to a commander. The video will be of a quality that is expected through existing and near future information systems. The purpose of this experiment is to try and determine the minimum bandwidth required for the observer to maintain their spatial perception. To help determine this you will be asked to perform two tasks.

Primary: Using a pen, as you watch the video you will be asked to plot your location in the environment on the floor plan provided at a set time interval (15 seconds). Also indicate with the tail of the check mark the direction you feel you are looking. The video will run continuously for a period of one and a half minutes.

Secondary: Please look at the objects depicted in the frame captures at the top if the board. These objects are from left to right, a floor space heater, a laser printer, a safe, a fire extinguisher, and a drinking fountain. Immediately after you have finished viewing the video I would like you to place these objects on the floor plan where you feel you saw them during the video using the colored arrow as provided.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

4I2I Corporation, [http://www.4i2i.com/], software and hardware implementations of H.263

Chairman of the Joint Chiefs of Staff, *Joint Vision 2020*, Pentagon, Washington, D.C., 2000

Headquarters, Department of the Army, *Field Manual 5-170 Engineer Reconnaissance*, Washington, D.C., 1998

Headquarters, Department of the Navy, United States Marine Corps, *Marine Corps Doctrinal Publication 6, Command and Control*, Washington, D.C., 1996

ITU-T Recommendation H.263, "Video coding for low bit rate communication"

Laplante, P.A., Stoyenko, A. D., *Real-Time Imaging-Theory, Techniques and Applications*, IEEE Press, 1996

Microsoft Media Technologies Jumpstart CD, Dec 1999

Monroe, J.E., Passagian, A.J., *Digital Video Imagery and Wireless Communications For Landbased Reconnaissance Missions*, Masters Thesis, Naval Post Graduate School December 1998

MSDN Online Workshop [http://msdn.microsoft.com/workshop/windowsmedia], 2000

Pe'ruch, P.,Vercher,J-L, Gathier, G.M, "Acquisition of Spatial Knowledge Through Visual Exploration of Simulated Environments", *Ecological Pychology,* 1995

Raghavan, S.V., Tripathi, S.K., *Networked Multimedia Systems: Concepts, Architecture and Design*, Prentice Hall, 1997

Rao, K.R., Hwang, J.J., *Techniques & Standards for Image, Video and Audio Coding*, Prentice Hall PTR, 1996

Reese, Mark, RF Engineer, Raytheon Corporation, Phone Intereview March 13, 2000

Sharp Corporation, News Release, *Sharp Unveils World's First Digital Video Camera with Advanced MPEG-4 Compression*, [http://www..sharp-usa.com/], June 22, 1999

Teledesic, [http://www.teledesic.com/tech/tech.htm], July 25, 2000

Wagonner, Ben, “Choosing a Streaming Video Technology”, *DV Web Video*, May 2000

Zimm, Alan D., CDR, U.S. Navy (Retired). "Human-Centric Warfare," *Proceedings*, pp. 28-31, May 1999

INTIAL DISTRIBUTION LIST

1. Defense Technical Information Center (2)
   8725 John J. Kingman Road, Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library (2)
   Naval Postgraduate School
   411 Dyer Road
   Monterey, CA 93943-5101

3. Director, Training and Education (1)
   MCCDC, Code C46
   1019 Elliot Road
   Quantico, VA 22134-5027

4. Director, Marine Corps Research Center (2)
   MCCDC, Code C40RC
   2040 Broadway Street
   Quantico, VA 22134-5107

5. Marine Corps Representative (1)
   Naval Postgraduate School
   Code 037, Bldg. 330, Ingersoll Hall, Room 116
   555 Dyer Road
   Monterey, CA 93943

6. Marine Corps Tactical Systems Support Activity (1)
   Technical Advisory Board
   Attn: Librarian
   Box 555171
   Camp Pendleton, CA 92055-5080

7. Lois B. Kempster (1)
   1145 North Shannon Ave. #4
   Indialantic, FL 32903

8. Kurt Kempster (1)
   229 Anzio Rd.
   Seaside, CA 93955

9. Rudolph P. Darken (1)
Naval Postgraduate School
Code CS/DR
Monterey, CA 93943